



QY 50 ATACAGCAATGTGAAACACGAGGTCATGCTGAGGTTTCATCAAAACACTGGAAGCCA 109  
 |||||  
 Db 335 ATAGTACTGTTGGCCAGAGGCCATGCACTGAGGTTTTCCTCAACACTGAAGCCC 394  
 |||||  
 QY 110 TGGTAGTATTGTCCTTGTAGGTATGTCACCTCAACCAATGTCCTTCTGGCTGAT 169  
 |||||  
 Db 395 TGGTAGTATTGTCCTTGTAGGTATGTCACCTCAACCAATGTCCTTCTGGCTGAT 454  
 |||||  
 QY 170 GGAGAAATGGGCTGC 184  
 |||||  
 Db 455 GGAGAAATGGGCTGC 469  
 |||||

RESULT 2  
 H94042/c  
 LOCUS  
 DEFINITION  
 Yv15d06.8 Soares fetal liver spleen INFLS Homo sapiens cDNA clone  
 IMAGE:242795 3' similar to gb:X13238 CYTOCHROME C OXIDASE  
 POLYPEPTIDE VIC PRECURSOR (HUMAN); mRNA sequence.

ACCESSION  
 VERSION  
 H94042  
 H94042.1 GI:1101338  
 EST.  
 EST.  
 Homo sapiens (human)  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homnidae; Homo.  
 1 (bases 1 to 419)  
 Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,  
 Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,  
 Parsons, J., Rifkin, L., Rohlffing, T., Soares, M., Tan, F.,  
 Trevaekis, E., Waterston, R., Williamson, A., Wohlmann, P. and  
 Wilson, R.  
 The WashU-Merck EST Project  
 Unpublished (1995)  
 Contact: Wilson RK  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 High quality sequence stops: 277  
 Source: IMAGE Consortium, LLNL  
 This clone is available royalty-free through LLNL; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 Insert Length: 616 Std Error: 0.00  
 Seq primer: m13 -40 forward.

FEATURES  
 source  
 1..419  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="GDB:3791928"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:242795"  
 /sex="male"  
 /dev\_stage="20 week-post conception fetus"  
 /lab\_host="DH10B (ampicillin resistant)"  
 /clone\_lib="Soares fetal liver spleen INFLS"  
 /note="Organ: Liver and Spleen; Vector: p7T3D (Pharmacia)  
 with a modified polylinker; Site 1: Pac I; Site 2: Eco RI;  
 1st strand cDNA was primed with a Pac I - oligo(dT) primer  
 [5' AACTGGAAGATTAATTAAGATCTTTTCTTTTCTTTT 3'],  
 double-stranded cDNA was ligated to Eco RI adaptors  
 (Pharmacia), digested with Pac I and cloned into the Pac I  
 and Eco RI sites of the modified p7T3 vector. Library  
 went through one round of normalization. Library  
 constructed by Bento Soares and M.Fatima Bonaudo."

ORIGIN  
 Query Match 24.2%; Score 89.6; DB 8; Length 419;  
 Best Local Similarity 68.8%; Pred. No. 6.3e-12;  
 Matches 137; Conservative 0; Mismatches 60; Indels 2; Gaps 1;

QY 18 ATGTATAGTGTCTCAACCCCTCTATACAGGATACAGCAATGTGAAACACGAGGTCATG 77  
 |||||  
 Db 266 AGGATACAGTGAATGCTCCCAATATGAAATGCGAATGGTGCAGAGGCCACATG 207  
 |||||  
 QY 78 CATCTGAGGTTTCATCAAAACACTGGAAGCCATGAGTATGTTCTTCTGATAGGTATGT 137  
 |||||  
 Db 206 --TCTGAGGTTTCATGAAAGCACTGGAACACCATGGTAGTTCTGTTCTTTGATAGTATGT 149  
 |||||  
 QY 138 CAACCTCAACACCAATGTCCTTCTGGCTGATGGAGAAATGGCTGCCGATTCACAGAC 197  
 |||||  
 Db 148 CGGCTCAACACCAAGTCCCTTCTGGCTGATGAAGAAATGGATTCACCTTCCCTGAA 89  
 |||||  
 QY 198 TTCTTGATCTTCTCTCAGC 216  
 |||||  
 Db 88 AGGATGTCAGTGCCTCAGC 70  
 |||||

RESULT 3  
 T91795/c  
 LOCUS  
 DEFINITION  
 Yd52h12.8 Soares fetal liver spleen INFLS Homo sapiens cDNA clone  
 IMAGE:114911 3' similar to gb:X13238 CYTOCHROME C OXIDASE  
 POLYPEPTIDE VIC PRECURSOR (HUMAN); mRNA sequence.

ACCESSION  
 VERSION  
 T91795  
 T91795.1 GI:723708  
 EST.  
 EST.  
 Homo sapiens (human)  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homnidae; Homo.  
 1 (bases 1 to 456)  
 Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,  
 Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,  
 Parsons, J., Rifkin, L., Rohlffing, T., Soares, M., Tan, F.,  
 Trevaekis, E., Waterston, R., Williamson, A., Wohlmann, P. and  
 Wilson, R.  
 The WashU-Merck EST Project  
 Unpublished (1995)  
 Contact: Wilson RK  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 Insert Size: 614  
 High quality sequence stops: 272 Source: IMAGE Consortium, LLNL  
 This clone is available royalty-free through LLNL; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 Insert Length: 614 Std Error: 0.00  
 Seq primer: -21m13  
 High quality sequence stop: 272.

FEATURES  
 source  
 1..456  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="GDB:467528"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:111911"  
 /sex="male"  
 /dev\_stage="20 week-post conception fetus"  
 /lab\_host="DH10B (ampicillin resistant)"  
 /clone\_lib="Soares fetal liver spleen INFLS"  
 /note="Organ: Liver and Spleen; Vector: p7T3D (Pharmacia)  
 with a modified polylinker; Site 1: Pac I; Site 2: Eco RI;  
 1st strand cDNA was primed with a Pac I - oligo(dT) primer  
 [5' AACTGGAAGATTAATTAAGATCTTTTCTTTTCTTTT 3'],  
 double-stranded cDNA was ligated to Eco RI adaptors  
 (Pharmacia), digested with Pac I and cloned into the Pac I  
 and Eco RI sites of the modified p7T3 vector. Library  
 went through one round of normalization. Library  
 constructed by Bento Soares and M.Fatima Bonaudo."

ORIGIN



```

Query Match      24.2%; Score 89.6; DB 8; Length 456;
Best Local Similarity 67.2%; Pred. No. 6.3e-12;
Matches 125; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 31 TACAAACCTCTATACAGGATGATGAGCAATGTAACACGAGGTCATGATCTGAGGTTTC 90
DB 236 TGCTCCCCACAATATGAATTTGCCAGANTGTTGCCAGAGGCCACACATGTCAGGTTTC 177

QY 91 ATCAAAACACCTGGAGCCATGTTAGTTATTCCTTGTATAGTAGTATGTCACACTCAACACC 150
DB 176 ATGAAGACCTGGAACCATGTTAGTTACTGTTCTTGATAGTAGTATGTCGGCTCAACACC 117

QY 151 AATGTCCTTCTCGCTGATGAGAAATGGGCTGCCGATTCACAGACTTCTTGATCTTTC 210
DB 116 AGTGCCCTTCTCGCTGATGAGAAATGGGCTGCCGATTCACAGACTTCTTGATCTTTC 210

QY 211 CTCAGC 216
DB 56 CCCAGC 51

RESULT 4
AV734452/c
LOCUS
DEFINITION AV734452 cda Homo sapiens cDNA clone cdaAOA01 5', mRNA sequence.
ACCESSION AV734452
VERSION AV734452.1 GI:10851997
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 456)
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
FEATURES
Location/Qualifiers
1..456
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="cdaAOA01"
/tissue_type="pneocromocytoma"
/dev_stage="Adult"
/lab_host="BM25.8"
/clone_lib="cda"
/note="Vector: pTriplEx2; Site_1: sf11A; Site_2: sf11B"

ORIGIN
Query Match      23.6%; Score 87.4; DB 1; Length 456;
Best Local Similarity 69.8%; Pred. No. 2.3e-11;
Matches 118; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 202 TGATCTTTCTCAGCATATGATCTCTCATATGAAACTCTCTCTGCTAGGAGTGCA 261
DB 399 TAATATCTTTAAGCATGACATAAGTAATGAATTAAGCATGCTAGCATGTA 340

QY 262 GCCTACATCGCTAAATCCAGCACCTTGGGAAGCTGAGGAGGAGGATTAATCGAGCCAG 321
DB 339 GCTCACATCTGCAATCCAGCACCTTGGGAGGCTGAGGAGGATGACTGCTTGGAGCCAG 280

```

```

QY 322 GAGTTTGAGACCGGCTGGGCAATACAGCGAGACTCTCTCTCTTAAAAA 370
DB 279 GAGTTTCAAGACCGGCTGGGCAACATAGTGAGACCCCATCTCCAAAAA 231

RESULT 5
BG912371/c
LOCUS
DEFINITION BG912371 NCI CGAP_Brn67 Homo sapiens cDNA clone IMAGE:4940673
5', mRNA sequence.
ACCESSION BG912371
VERSION BG912371.1 GI:14292847
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 744)
Email: cgaps-remail.nih.gov
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
JOURNAL
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: David N. Louis, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10879 row: k column: 10
High quality sequence stop: 711.
FEATURES
Location/Qualifiers
1..744
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4940673"
/tissue_type="anaplastic oligodendroglioma with 1p/19q
loss"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI CGAP Brn67"
/note="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2.3 kb. Constructed by Life
Technologies. Note: this is a NCI CGAP Library."

ORIGIN
Query Match      23.4%; Score 86.6; DB 2; Length 744;
Best Local Similarity 80.8%; Pred. No. 3.7e-11;
Matches 101; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 246 CTGGCTAGGAGTGGCAGCCTACATCGCTAATCCAGCACCTTGGGAAGCTGAGGAGAG 305
DB 485 CTGGCTAGGCGTGGTGGCCACATAAGTAATACAGCAGCTTTGGGAGGCCAAGTCAGGAG 426

QY 306 GATTACTGGAGGCCAGGAGTTTGAGACCGAGCTGGGCAATACAGCGAGACTCTCTCT 365
DB 425 GATCATTGAGGCCAGGAGTTTGAGACCAAGCCTGGGCAACATAGTGAGACTCCATCTCT 366

QY 366 AAAAA 370
DB 365 AAAAA 361

RESULT 6
AUI56956/c
LOCUS
DEFINITION AUI56956 PLACE1 Homo sapiens cDNA clone PLACE1005453 3', mRNA
sequence.
ACCESSION AUI56956

```

**VERSION** AUI56956.1 GI:11018477  
**KEYWORDS** EST.  
**SOURCE** Homo sapiens (human)  
**ORGANISM** Homo sapiens  
**REFERENCE** 1 (bases 1 to 568)  
**AUTHORS** Oca,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y., Yamamoto,J., Wakamatsu,A., Ozawa,M., Nakamura,Y., Nagai,T., Sugano,S. and Isogai,T.).  
**TITLE** HRI human cDNA project (Oca,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y., Yamamoto,J., Wakamatsu,A., Ozawa,M., Nakamura,Y., Nagai,T., Sugano,S., Isogai,T.)  
**JOURNAL** Unpublished (2000)  
**COMMENT** Contact: Takao Isogai  
 Genomics Laboratory  
 Helix Research Institute  
 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
 Tel: 81-438-52-3975  
 Fax: 81-438-52-3986  
 Email: genomics@hri.co.jp  
 HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.  
**FEATURES** Location/Qualifiers  
 source 1..568  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="PLACE1005453"  
 /tissue\_type="placenta"  
 /clone\_lib="PLACE1"  
 /note="Vector: pME18SFL3"  
**ORIGIN**  
 Query Match 23.4%; Score 86.4; DB 1; Length 568;  
 Best Local Similarity 68.8%; Pred. No. 4.1e-11;  
 Matches 117; Conservative 0; Mismatches 53; Indels 0; Gaps 0;  
 QY 201 TTGATCTTCTCAGCATATGATCTCTCATATGAACTCTCTCTGGCTAGGATGGC 260  
 DB 331 TTTATTTATTTATCAATACATATAGTGAAGTGGATCTCTGGCTAGTGCAGT 272  
 QY 261 AGCTACATCTGAATCCAGCACCTTTGGGAAGCTGAGGAGGAGGATTAATCGAGGCCA 320  
 DB 271 GGCTCATGCTGTAACTCTAGCACTTTGGGAGGCTGAGGAGGAGGATTCCTTGAGCCCA 212  
 QY 321 GGAGTTTGAGACCAAGCTGGGCAATACAGCGAGACTCTCTCTCTAAAAA 370  
 DB 211 GGAGTCTGAGACCAAGCTGGGCAACATAGGAGACTCTGTCTGCAAAA 162  
**RESULT 7**  
**LOCUS** BM755930  
**DEFINITION** K-EST0034028 SISNUS Homo sapiens cDNA clone SISNUS-35-B01 5', mRNA sequence.  
**ACCESSION** BM755930  
**VERSION** BM755930.1 GI:19085545  
**KEYWORDS** EST.  
**SOURCE** Homo sapiens (human)  
**ORGANISM** Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.  
**REFERENCE** 1 (bases 1 to 601)  
**AUTHORS** Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.J., Cheong,J.B., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S.  
**TITLE** 21C Frontier Korean EST Project 2001  
**JOURNAL** Unpublished (2002)  
**COMMENT** Contact: Kim YS  
 Genome Research Center  
 Korea Research Institute of Bioscience & Biotechnology  
 52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
 Tel: +82-42-860-4470  
 Fax: +82-42-860-4409  
 Email: yongsung@mail.kribb.re.kr  
 Plate: 35 row: B column: 01  
 High quality sequence stop: 601.  
 Location/Qualifiers  
 source 1..601  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="SISNUS-35-B01"  
 /sex="F"  
 /tissue\_type="Ascites"  
 /cell\_type="Lymphoblast-like"  
 /cell\_line="SNU-5"  
 /lab\_host="Top10F"  
 /clone\_lib="SISNUS"  
 /note="Organ: Stomach; Vector: pCNS; Site:1: EcoRI; Site 2: NotI; The poly (A) + RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tabacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including EcoR I site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dt-selected mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."  
**ORIGIN**  
 Query Match 23.2%; Score 86; DB 3; Length 601;  
 Best Local Similarity 71.5%; Pred. No. 5.2e-11;  
 Matches 113; Conservative 0; Mismatches 45; Indels 0; Gaps 0;  
 QY 213 CAGCATATGATCTCTCATATGAACTCTCTCTGGCTAGGATGGCAGCTACATCGC 272  
 DB 422 CAGGCGAAGATGCATCTCTATAAAGACTCAAGAAGTCCAGGTGTGGTGTCTCACACCTG 481  
 QY 273 TAATCCAGCACTCTGGGAAGCTGAGGAGGAGGATTAATCTGGAGGCGAGGATTTGAGAC 332  
 DB 482 TAATCCAGCACTTTGGGAAGGCGAGGAGGATCCCTTGAGGCGAGGATTCGAGAC 541  
 QY 333 CAGGCTGGGCAATACAGCGAGACTCTCTCTCTCTAAAAA 370  
 DB 542 CAGCCTGGGCAACATAGCAAGACTCCGTCTCTCAAAAAA 579  
**RESULT 8**  
**LOCUS** BF211330/c  
**DEFINITION** 601812572F1 NIH\_MGC\_54 Homo sapiens cDNA clone IMAGE:4047209 5', mRNA sequence.  
**ACCESSION** BF211330  
**VERSION** BF211330.1 GI:11104916  
**KEYWORDS** EST.  
**SOURCE** Homo sapiens (human)  
**ORGANISM** Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.  
**REFERENCE** 1 (bases 1 to 673)  
**AUTHORS** NIH-MGC http://imgc.nci.nih.gov/.  
**TITLE** National Institutes of Health, Mammalian Gene Collection (MGC)  
**JOURNAL** Unpublished (1999)

```
COMMENT      Contact: Robert Strausberg, Ph.D.
             Email: cgapbs-rc@mail.nih.gov
             Tissue Procurement: ATCC
             cDNA Library Preparation: CLONETECH Laboratories, Inc.
             DNA sequencing by: The I.M.A.G.E. Consortium (LNL)
             cDNA distribution: MGC clone distribution information can be
             found through the I.M.A.G.E. Consortium/LNL at:
             http://image.lnl.gov
             Plate: L1CM875 row: 0 column: 18
             High quality sequence stop: 560.

FEATURES     Location/Qualifiers
source       1..673
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone_lib="NIH_MGC_54"
             /note="Organ: bone marrow; Vector: pDNR-LIB (Clontech);
             Site 1: SfiI (ggcgcctggcc); Site 2: SfiI
             (ggcattatggcc); Double-stranded cDNA was prepared from
             cell line RNA. 5' and 3' adaptors were used in cloning as
             follows: 5' adaptor sequence: 5'-CACGGCCATTATGGCC-3' and
             3' adaptor sequence:
             5'-ATTCTAGAGCGGCGGCGGCGCATG-dT(30)BN-3' (where B = A,
             C, or G and N = A, C, G, or T). Average insert size
             1.75 kb (range 0.9-4.0 kb). 15/15 colonies contained
             inserts by PCR. This library was enriched for full-length
             clones and was constructed by Clontech Laboratories (Palo
             Alto, CA)."

ORIGIN
Query Match      23.2%; Score 86; DB 2; Length 673;
Best Local Similarity 80.2%; Pred. No. 5.2e-11;
Matches 101; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

Qy 245 CCTGGCTAGGAGTGGCAGCTTACATCGCTATCCAGACCTTGGGAAGCTGAGGCGAGGA 304
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 256 CCTGGCTGGGCGTGATCTCATGCTGAAATCCAGCACTTTGAGAGGCTTGACACAGGA 197
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 305 GGATTACTGGAGGCGCAGAGTTTGAGACCGGCTGGGCAATACAGCGAGACTCTCTCC 364
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 196 GGATTGTAGAGCGCAGAGTTTGAGACCGACCTGGTCAACACAGCGAGACCCCATCTCT 137
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 365 TAAAAA 370
    |||||
Db 136 AAAAAA 131

RESULT 9
AZ695087
LOCUS
DEFINITION      462 bp DNA linear GSS 31-JAN-2001
                 UP 501-13C SP6 RPC111 Human Male BAC Library Homo sapiens genomic
                 clone 501-13C, genomic survey sequence.
ACCESSION      AZ695087
VERSION        AZ695087.1 GI:12331723
KEYWORDS       GSS.
SOURCE         Homo sapiens (human)
ORGANISM       Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 462)
Chung,V.G., Dalrymple,H.L., Narasimhan,S., Watts,J., Schuler,G.,
Raap,A.K., Morley,M. and Bruzel,A.
A resource of mapped human bacterial artificial chromosome clones
Genome Res. 9 (10), 989-993 (1999)
10523527
PUBMED
COMMENT        Contact: Arcaro MA, Morley M, Burdick J, Cheung VG
                 Department of Pediatrics
                 University of Pennsylvania

FEATURES     Location/Qualifiers
source       1..590
             /organism="Homo sapiens"
             /mol_type="genomic DNA"
             /db_xref="taxon:9606"
             /clone="MCF7_1-14124"
             /sex="female"

3516 Civic Center Blvd, ARC 516, Philadelphia, PA 19104, USA
Tel: 215 590 2664
Fax: 215 590 3709
Email: mlennox@mail.med.upenn.edu
Plate: 501 row: 0 column: 13
Seq primer: SP6
Class: BAC ends.
Location/Qualifiers
1..462
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="501-13C"
/sex="Male"
/cell_type="Lymphocytes"
/clone_lib="RPC111 Human Male BAC Library"
/note="Vector: pBACe3.6; RPC111 Human Male BAC Library"

FEATURES     Location/Qualifiers
source       1..462
             /organism="Homo sapiens"
             /mol_type="genomic DNA"
             /db_xref="taxon:9606"
             /clone="501-13C"
             /sex="Male"
             /cell_type="Lymphocytes"
             /clone_lib="RPC111 Human Male BAC Library"
             /note="Vector: pBACe3.6; RPC111 Human Male BAC Library"

ORIGIN
Query Match      23.2%; Score 85.8; DB 9; Length 462;
Best Local Similarity 72.5%; Pred. No. 5.8e-11;
Matches 111; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy 218 TATGATACTTCATATGAACCTCTCTCTGGCTAGGAGTGGCAGCCTACATCGCTAATC 277
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 274 TTTGAGATATGATATAATATCTTGATGCTTTGGCTAGGTGCAGTGGCTCACACCTGCAATC 333
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 278 CCAGCACCTTTGGGAAGCTGAGGCGAGGAGGATTAATCTGAGGCGCAGGAGTTTGAGACCCAGGC 337
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 334 CTAGCATTGTTGGAGGCGAGGAGGAGGATCACTTCTGAGGCCAAGAGTTTGAGACCCAGGC 393
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 338 TGGGCAATACAGCGAGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 470
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 394 TGGGCAACAGAGTGAAACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 426
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

BZ606248
WHACF60TR Human MCF7 breast cancer cell line library (MCF7.1) Homo
sapiens genomic clone MCF7_1-14124, genomic survey sequence.
BZ606248
GSS.
BZ606248.1 GI:31514809
SOURCE         Homo sapiens (human)
ORGANISM       Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 590)
Volik,S., Zhao,S., Chin,K., Brebner,J.H., Herndon,D.R., Tao,Q.,
Kowbel,D., Huang,G., Lapuk,A., Kuo,W.-L., Magrane,G., de Jong,P.,
Gray,J.W. and Collins,C.
End-sequence profiling: Sequence-based analysis of aberrant genomes
Proc. Natl. Acad. Sci. U.S.A. 100 (13), 7696-7701 (2003)
12788976
PUBMED
COMMENT        Contact: Volik SV
                 Colin Collins' lab
                 UCSF Comprehensive Cancer Center
                 UCSF Box 0808, San Francisco, CA 94143-0808, USA
                 Tel: 415 502 7066
                 Fax: 415 502 5665
                 Email: svolik@cc.ucsf.edu
                 This clone is available from Amplicon Express
                 http://www.genomex.com
                 Class: BAC ends.
                 Location/Qualifiers
                 1..590
                 /organism="Homo sapiens"
                 /mol_type="genomic DNA"
                 /db_xref="taxon:9606"
                 /clone="MCF7_1-14124"
                 /sex="female"
```





Db 266 GC|TT|TT|TG|GCA|TAT|G|TT|CAG|TCC|GG|T|TAA|TTG|TAA|CT|AG|CT|GG|T|GAG|GTG 207  
Qy 263 CCTACATCGCTAATCCAGCACCTTGGGAAGCTGAGCAGGAGGATTACTGGAGGCCAGG 322  
Db 206 CTCACACCTGTAAATCCAGCGCTTTGGGAGGCTGAGGCGGTAGATTGTTTGAGCTCAGG 147  
Qy 323 AGTTTGAGACGAGCTGGGCAATACAGCAGAGACTCTCTCTCCTAA|AAA 370  
Db 146 GGTTCAGACAGCCTGCGCAATAAAGTGAGACTCTGTCTCTCA|AAA 99

## RESULT 15

B92454/c

LOCUS

DEFINITION B92454 459 bp DNA linear GSS 25-JUN-1998  
CIT-HSP-2172114.TF CIT-HSP Homo sapiens genomic clone 2172114,  
genomic survey sequence.

ACCESSION

B92454

VERSION

B92454.1 GI:2974791

KEYWORDS

GSS.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominoidea; Homo.1. (bases 1 to 459)  
Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,  
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,  
Simon,M. and Venter,J.C.  
Use of a random BAC End Sequence Database for Sequence-Ready Map  
Building (1998)

REFERENCE

Unpublished (1998)

JOURNAL

Other\_GSSs: CIT-HSP-2172114.TR

COMMENT

Contact: Mark Adams  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: mdadams@tigr.org  
Clones are available from Research Genetics (info@resgen.com). BAC  
end search page:  
[http://www.tigr.org/tdb/humgen/bac\\_end\\_search/bac\\_end\\_search.html](http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html)  
Seq primer: M13-21;  
Class: BAC ends.

## FEATURES

Source

Location/Qualifiers

1. .459  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="GDB:7103880"  
/db\_xref="taxon:9606"  
/clone="2172114"  
/sex="Male"  
/cell\_type="Sperm"  
/clone\_lib="CIT-HSP"  
/notes="Vector: pBelosBAC11; Site\_1: HindIII; Site\_2:  
HindIII"

## ORIGIN

Query Match 22.9%; Score 84.6; DB 9; Length 459;  
Best Local Similarity 69.0%; Pred. No. 1e-10;  
Matches 116; Conservative 0; Mismatches 52; Indels 0; Gaps 0;  
Qy 203 GATCTTTCTCAGCATATGATCTCTCATATGAAACTCTCTCTCTGGCTAGGAGTGGCAG 262  
Db 275 GC|TT|TT|TG|GCA|TAT|G|TT|CAG|TCC|GG|T|TAA|TTG|TAA|CT|AG|CT|GG|T|GAG|GTG 216  
Qy 263 CCTACATCGCTAATCCAGCACCTTGGGAAGCTGAGCAGGAGGATTACTGGAGGCCAGG 322  
Db 215 CTCACACCTGTAAATCCAGCGCTTTGGGAGGCTGAGGCGGTAGATTGTTTGAGCTCAGG 156  
Qy 323 AGTTTGAGACGAGCTGGGCAATACAGCAGAGACTCTCTCTCCTAA|AAA 370

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 21:43:03 ; Search time 751.729 Seconds  
(without alignments)  
6427.437 Million cell updates/sec

Title: US-09-980-046B-10  
Perfect score: 85  
Sequence: 1 cggtagtgatctcttcgc.....tgcccactaagtagagaaaa 85

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5983141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_in.\*

3: gb\_env.\*

4: gb\_on.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pr.\*

9: gb\_ro.\*

10: gb\_sts.\*

11: gb\_sy.\*

12: gb\_un.\*

13: gb\_vi.\*

14: gb\_htg.\*

15: gb\_pl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	83.4	98.1	451	6	AX135336 Sequence
2	83.4	98.1	1318	8	BC001114 Homo sapi
3	83.4	98.1	3901	8	HUMLP7A
4	83.4	98.1	19936	8	EX682530 Human DNA
5	83.4	98.1	19937	8	CR749765 Human DNA
6	83.4	98.1	35167	8	CT009502 Human DNA
7	83.4	98.1	47546	8	CR753889 Human DNA
8	83.4	98.1	63179	8	AL935043 Human DNA
9	83.4	98.1	66109	6	AX330551
10	83.4	98.1	66109	6	AX411124
11	83.4	98.1	66109	8	HSMHCAPG
12	83.4	98.1	95554	8	CR762476 Human DNA
13	83.4	98.1	110856	8	EX927138 Human DNA
14	83.4	98.1	114576	8	EX088556 Human DNA
15	83.4	98.1	116859	14	CR788563
16	83.4	98.1	126988	8	AL669918 Human DNA
17	83.4	98.1	166803	8	AL671681 Human DNA
18	83.4	98.1	177571	14	AC022567 Homo sapi

19	83.4	98.1	198285	6	CQ861705
20	83.4	98.1	198285	6	AX411170
21	83.4	98.1	198285	6	HSEVMHC
22	83.4	98.1	349980	6	CS039424
23	82.4	96.9	9362	8	AY559263
24	82.4	96.9	9362	8	AY559271
25	82.4	96.9	9362	8	AY559277
26	82.4	96.9	9362	8	AY559281
27	82.4	96.9	9362	8	AY559293
28	82.4	96.9	9362	8	AY559299
29	82.4	96.9	9363	8	AY559261
30	82.4	96.9	9366	8	AY559253
31	82.4	96.9	9366	8	AY559257
32	82.4	96.9	9366	8	AY559259
33	82.4	96.9	9366	8	AY559264
34	82.4	96.9	9366	8	AY559267
35	82.4	96.9	9366	8	AY559275
36	82.4	96.9	9366	8	AY559279
37	82.4	96.9	9366	8	AY559280
38	82.4	96.9	9366	8	AY559281
39	82.4	96.9	9366	8	AY559285
40	82.4	96.9	9366	8	AY559295
41	82.4	96.9	9366	8	AY559297
42	82.4	96.9	9367	8	AY559273
43	82.4	96.9	9367	8	AY559289
44	82.4	96.9	9368	8	AY559260
45	82.4	96.9	9368	8	AY559262

ALIGNMENTS

RESULT 1  
AX135336  
LOCUS AX135336 451 bp DNA linear PAT 29-MAY-2001  
DEFINITION Sequence 144 from Patent WO0132928.  
ACCESSION AX135336  
VERSION AX135336.1 GI:14271685  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE 1  
AUTHORS Farr S.  
TITLE Methods of determining individual hypersensitivity to an agent  
JOURNAL Patent: WO 0132928-A 144 10-MAY-2001;  
FEATURES Phase-1 Molecular Toxicology Inc. (US)  
Location/Qualifiers  
source 1. 451  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

Query Match 98.1%; Score 83.4; DB 6; Length 451;  
Best Local Similarity 98.8%; Pred. No. 5.8e-18;  
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 CGGTGAAGTGATCTTCTCGGTGTTCTCTATTGAACAGCATTTCCCGGAGGAGTTT 60  
Db 353 CGGTGAAGTGATCTTCTCGGTGTTCTCTATTGAACAGCATTTCCCGGAGGAGTTT 412  
Qy 61 CTGGGTGCCCCACTAAGTAGAGAAA 85  
Db 413 CTGGGTGCCCCACTAAGTAGAGATAA 437

RESULT 2  
BC001114  
LOCUS BC001114 1318 bp mRNA linear PRI 28-JUL-2005  
DEFINITION Homo sapiens proteasome (prosome, macropain) subunit, beta type, 8

(large multifunctional protease 7), transcript variant 1, mRNA  
(CDNA clone MGC:1491 IMAGE:2967119), complete cds.

BC001114  
MGC.  
BC001114.2 GI:33876008  
MGC.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominiidae; Homo.  
1 (bases 1 to 1318)  
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,  
Klausner, R.D., Collins, F.S., Wagner, L., Shennan, C.M., Schuler, G.D.,  
Altschul, S.F., Zeeberg, B., Buettow, K.H., Schaefer, C.F., Bhat, N.K.,  
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heide, F.,  
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,  
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,  
Scheetz, T.E., Brownstein, M.J., Ustin, T.B., Toshiyuki, S.,  
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,  
Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,  
McKernan, R.J., Malek, J.A., Gunaratne, P.H., Richards, S.,  
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,  
Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,  
Fahey, J., Helton, E., Kettelman, M., Madan, A.C., Rodriguez, S.,  
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,  
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,  
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,  
Butterfield, Y.S., Krzyzaniak, M.I., Skalska, U., Smalios, D.E.,  
Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.  
Mammalian Gene Collection Program Team  
Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
12477932  
2 (bases 1 to 1318)  
NITH MGC Project  
Direct Submission  
Submitted (11-DEC-2000) National Institutes of Health, Mammalian  
Gene Collection (MGC), Bethesda, MD 20892-2590, USA  
NITH-MGC Project URL: <http://mgc.nci.nih.gov>  
On Aug 19, 2003 this sequence version replaced gi:12654558.  
Contact: MGC help desk  
Email: [cgap@nci.nih.gov](mailto:cgap@nci.nih.gov)  
Tissue Procurement: ATCC/DCTB/DTF  
CDNA Library Preparation: Ruben Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Institute for Systems Biology  
<http://www.systemsbio.org>  
contact: [amadan@systemsbio.org](mailto:amadan@systemsbio.org)  
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha  
Madan, Stephanie Rodriguez, Amy Sanchez and Michelle Whiting  
Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAL Plate: 3 Row: m Column: 19  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA gi: 34335277.  
Location/Qualifiers  
1. .1318  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="MGC:1491 IMAGE:2967119"  
/tissue\_type="Skin, melanotic melanoma."  
/clone\_lib="NITH MGC 20"  
/lab\_host="DH10B-R"  
/notes="vector: pOTB7"  
1. .1318  
/genes="PSMB8"  
/notes="synonyms: RING10, MGC1491, D6S216, D6S216E"  
/db\_xref="GeneID:5696"  
/db\_xref="MIM:177046"

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
PUBMED  
AUTHORS  
CONSTRM  
TITLE  
JOURNAL  
PUBMED  
AUTHORS  
CONSTRM  
TITLE  
REMARK  
COMMENT  
FEATURES  
source  
gene

1057  
/genes="PSMB8"  
/codon\_start=1  
/product="proteasome beta 8 subunit, isoform E1  
proteasome"  
/protein\_id="AAH01114.1"  
/db\_xref="GI:12654559"  
/db\_xref="GeneID:5696"  
/db\_xref="MIM:177046"  
/translation="MLIGTTPRDTTPSSWLTSLLVEAAPLDDTTLTPVSSGCGPL  
EPTFQSLGSGDGRNVOIEMAHGTTTLLAFKQHGVIIVAAVDSRASAGSYISARVNVK  
IINPVLCTGMSGCAADCOYWERLLAKELRYLRNGERISVSAASKLLSNMCMOYRG  
MELSGSMICGDKKPGIYYDEHGTRELRLSGMFGSTGNTYAYGWMDSGYRNLSP  
EAYDLGRRAIATATHRDSYSGGVNMIHMKDGVKVESTVDSDLLHQIREANQ"

ORIGIN  
Query Match 98.1%; Score 83.4; DB 8; Length 1318;  
Best Local Similarity 98.8%; Pred. No. 5.8e-18;  
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGTAACAAGCATTTCCCCAGGGAAGTTT 60  
Db 1196 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGTAACAAGCATTTCCCCAGGGAAGTTT 1255  
QY 61 CTGGTGCCCCACTAAGTAGAAAA 85  
Db 1255 CTGGTGCCCCACTAAGTAGAAAA 1280  
RESULT 3  
HUMIMP7A HUMIMP7A  
LOCUS Human LMP7 gene, exons 1-7.  
DEFINITION Human LMP7 gene, exons 1-7.  
ACCESSION L11045  
VERSION L11045.1  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominiidae; Homo.  
REFERENCE  
AUTHORS Meinhardt, T., Graf, U. and Hammerling, G.J.  
TITLE Different genomic structure of mouse and human Lmp7 genes:  
characterization of MHC-encoded proteasome genes  
JOURNAL Immunogenetics 38 (5), 373-379 (1993)  
PUBMED 8344725  
COMMENT Original source text: Homo sapiens DNA.  
FEATURES  
source Location/Qualifiers  
1. .3901  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
/cell\_line="B lymphoblastoid cell line MANN"  
1. .3901  
/genes="LMP7"  
1. .3901  
/genes="LMP7"  
mRNA join(1. .355, 610. .784, 1545. .1692, 1851. .1962, 2368. .2497,  
2893. .3097, 3578. .3901)  
/genes="LMP7"  
1. .355  
/genes="LMP7"  
/number=1  
356. .609  
/genes="LMP7"  
/number=1  
610. .784  
/genes="LMP7"  
/number=2  
785. .1544  
/genes="LMP7"  
/number=2



1545. .1692  
/gene="LMP7"  
/number=3  
1693. .1850  
/gene="LMP7"  
/number=3  
1851. .1962  
/gene="LMP7"  
/number=4  
1963. .2367  
/gene="LMP7"  
/number=4  
2368. .2497  
/gene="LMP7"  
/number=5  
2498. .2892  
/gene="LMP7"  
/number=5  
2893. .3097  
/gene="LMP7"  
/number=6  
3098. .3577  
/gene="LMP7"  
/number=6  
3578. .3901  
/gene="LMP7"  
/number=7  
3885. .3890  
/gene="LMP7"

exon  
intron  
exon  
intron  
exon  
intron  
exon  
intron  
exon  
polya\_site  
ORIGIN

Query Match 98.1%; Score 83.4; DB 8; Length 3901;  
Best Local Similarity 98.8%; Pred. No. 5.7e-18;  
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGTGAAGTCATCTTCGCGTGTCTCTATTGAACCAAGCATTTCCCCAGGGAAGTTT 60  
Db 3805 CGGTGAAGTCATCTTCGCGTGTCTCTATTGAACCAAGCATTTCCCCAGGGAAGTTT 3864

Qy 61 CTGGGTGCCCACTAAGTAGAATA 85  
Db 3865 CTGGGTGCCCACTAAGTAGAATA 3889

RESULT 4  
BX682530/c  
LOCUS Human DNA sequence from clone DASS-193J13 on chromosome 6, complete sequence.  
DEFINITION  
ACCESSION BX682530  
VERSION BX682530.4 GI:42475735  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 19936)  
Palmer.S.  
Direct Submission  
Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk  
Clone requests: clonerequest@sanger.ac.uk  
On Feb 7, 2004 this sequence version replaced gi:40067399.  
The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:  
Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at  
http://www.sanger.ac.uk/projects/C.elegans/wormpep This sequence was generated from part of bacterial clone configs constructed by the MHC Haplotype Consortium and collaborators. Further information can be found at  
http://www.sanger.ac.uk/HGP/Chr6/MHC

DASS-193J13 is from a DNA-arts SSTO human bac library VECTOR:  
pBelOBAC11  
----- Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: http://www.sanger.ac.uk  
Contact: vegas@sanger.ac.uk  
-----  
This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

FEATURES  
Location/Qualifiers  
1. .19936  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="6"  
/clone="DASS-193J13"  
/clone\_lib="DNA-arts-BAC.1-SSTO.1"  
2000  
/note="Clone\_right\_end: DASS-25606"  
complement(join(7236..7759,8125..8261,8655..8814,8992..9165,9343..9531,11058..11186,11352..11549,13879..14084,14368..14498,16262..16376,16466..16962,17374..17465))  
/locus\_tag="DASS-193J13.1-001"  
complement(join(7236..7759,8125..8261,8655..8814,8992..9165,9343..9531,11058..11186,11352..11549,13879..14084,14368..14498,16262..16376,16466..16962,17374..17465))  
/locus\_tag="DASS-193J13.1-001"  
/notes="match: ESTs: Em:903732.1 Em:AUI32766.1 Em:BI84592.1 Em:BM561472.1 Em:BQ277877.1 match: CDNAS: Em:BC002751.2 Em:M74447.1 Em:M84748.1 Em:Z22936.1"  
complement(join(7580..7759,8125..8261,8655..8814,8992..9165,9343..9531,11058..11186,11352..11549,13879..14084,14368..14498,16262..16376,16466..16958))  
/locus\_tag="DASS-193J13.1-001"  
/standard\_name="OTTHUMP00000038912"  
/codon\_start=1  
/protein\_id="CAI41935.1"  
/db\_xref="GI:57209644"  
/db\_xref="GOA:Q5HY71"  
/db\_xref="InterPro:IPR001140"  
/db\_xref="InterPro:IPR003439"  
/db\_xref="InterPro:IPR003593"  
/db\_xref="InterPro:IPR005293"  
/db\_xref="UniProt/TREMBL:Q5HY71"  
/translation="MRPLDRLPWTSLLLVDAAALLLQGLTLLPGLPGLWLEGTLLRGVGLLKLGLLGFVGLTLLPLCLATPLTVSLRALVAGASRPAPARVASAPWSWL LVGGAAGLSLWAVLSPFCAEQDQVNNKVMRLKSLRDLPLVAAAFFLV LNLVGLTPIPHYSGRVIDILGDFDPAFAIAIFMCLFSFSGSLSGACGCGCTYTM SRINLRIRLOLFSSLLRQDFTFOETKTGELNSRLSDTTLMSNWLPLNANVLLRSLV KTVGLYGFMLSI SPLTLLSLHMPPTIAAEKVYNTRHQEVLRIOQAVARAGOVVRE AVGLQTVRFSGAEHEVCYKEALEOCRLQYWRDLERALLVLRVRLHGLVQMLM SCGLQMQQDGLTQGSLLSEMYQSSVSVYQTLVYIYGMLSLVGAAEKVFSIMDRQ KPLSPGTLPATTLQGVYKQDFSPAYPNRPDLVKGLTFLRPGVETALVAPNGSG KSTVAALLQNYLQPTGGVLLDEKPIQYEHCVLHSQVSVGQBPVLFSGSVRNIIAY GLOSCDDKYMAQAHAADDFTQEMEHGIYTDVGEKGSOLAAGQKORLALARLVDR PRVILDEATSLDVQCEQALQDWNNSRGDRVLIYVIAHLQAVQRAHQILVLEQGLKQ LAQEQGDLYSRVLVQQRLLMD"  
join(complement(17374..17465), complement(16466..16962), complement(16262..16376), complement(14368..14498), complement(13879..14084), complement(11352..11549), complement(11058..11186), complement(9343..9531), complement(8992..9165), complement(8655..8814), complement(8125..8261), complement(BX296564.7:73552..74037))

misc\_feature  
gene  
mRNA  
CDS  
gene

```

mRNA
/locus tag="DASS-193J13.1-002"
join(Complement(17374..17465),complement(16466..16962),
complement(16262..16376),complement(14368..14498),
complement(13879..14084),complement(11352..11549),
complement(11058..11186),complement(9343..9531),
complement(8992..9165),complement(8655..8814),
complement(8125..8261),
complement(BX296564.7:73552..74037))
/locus tag="DASS-193J13.1-002"
/notes="match: ESTs: Em:CD366061.1
match: cDNAs: Em:AF078671.1 Em:AF105151.1"
join(Complement(16466..16962),complement(16262..16376),
complement(14368..14498),complement(13879..14084),
complement(11352..11549),complement(11058..11186),
complement(9343..9531),complement(8992..9165),
complement(8655..8814),complement(8125..8261),
complement(BX296564.7:74008..74037))
/locus tag="DASS-193J13.1-002"
/standard name="OTTHUMP0000038914"
/codon start=1
/protein_id="CAI41936.1"
/db_xref="GI:57209645"
/db_xref="GOA:Q9UP03"
/db_xref="InterPro:IPR001140"
/db_xref="InterPro:IPR003439"
/db_xref="InterPro:IPR003593"
/db_xref="InterPro:IPR005293"
/db_xref="InterPro:IPR011527"
/db_xref="UniProt/TREMBL:Q9UP03"
/translations="MRLPDRPMTSLLLVDAALLWLLQGLPLTLLPQGLPLWLEGLT
RLGGLWLLKRLGLLGVLTLLPLCLATPLTVSLRALVAGASRAPPARVASAPWSML
LVGYGAAGLSMSLWAVLSPGAQOEKQDQVNNKVLWRLKLSRPDLPLIVAAFFLV
LAVLGTLPIHYSGRVIDILGDFDPHAFASAIFFWCLFSFGSSLSAGRCGGFTYTM
SRINLRERQLFSLLRQDLGFFQETKTGELNSRLSSDTTLMNSNWLPLNANVLLRSLV
KVYGLYFMLSISIPRTLTLSLHMPFTIAAEKYNTRHQEVLEIQQDAVARAQVYRE
AVGLQTVRSFGAEHEVRYKEALQCRQLYWRDLERALLYLVRVHLGVQMLML
SCGLQMQDELQGLLSLPMYQESVYQVTLVYIGDMLSNVGAEEKVFSYMDRQ
PNLPSPTGLAPTLQGVKFDQVSFAYPNRPDRPVLKGLTFTLRPEVTALVPGNSG
KSTVAALLQNYLPTQGVLLDKPKISOYHCVLHSQVVSQVGPQVLFSGSVRNINAY
GLQSCDDKVMNAAQAADDFIQEWEHGIYTDVGEKSQLAAGQKQRLAIARALVRD
PRVLIDEATSLADVDQCEQAKTLWKFMIF"
17208..17223
/misc_feature
17937
/notes="Single clone region"
/notes="Clone_left_end: DASS-67H14"

```

## ORIGIN

```

Query Match      98.1%; Score 83.4; DB 8; Length 19936;
Best Local Similarity 98.8%; Pred. No. 5.6e-18;
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 60
    |||||
Db 19543 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 19484
    |||||

QY 61 CTGGGTGCCCCACTAAGTAGAATAA 85
    |||||
Db 19483 CTGGGTGCCCCACTAAGTAGAATAA 19459
    |||||

```

## RESULT 5

```

CR749765/c
LOCUS
DEFINITION
Human DNA sequence from clone DASS-94G7 on chromosome 6, complete
sequence.
ACCESSION
CR749765
VERSION
CR749765.5 GI:56541856
KEYWORDS
HTG.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.

```

## REFERENCE

AUTHORS  
TITLE  
JOURNAL

## COMMENT

1 (bases 1 to 19937)  
Wood, J.  
Direct Submission  
Submitted (08-DEC-2004) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk  
On Dec 9, 2004 this sequence version replaced gi:56382936.  
----- Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: http://www.sanger.ac.uk  
Contact: humquery@sanger.ac.uk  
-----

During sequence assembly data is compared from overlapping clones.  
Where differences are found these are annotated as variations  
together with a note of the overlapping clone name. Note that the  
variation annotation may not be found in the sequence submission  
corresponding to the overlapping clone, as we submit sequences with  
only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all  
regions were either double-stranded or sequenced with an alternate  
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by at least  
one plasmid subclone or more than one M13 subclone; and the  
assembly was confirmed by restriction digest, except on the rare  
occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession  
numbers given in the feature table with their source databases:

Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information  
on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C\_elegans/wormpep This sequence  
was generated from part of bacterial clone contigs constructed by  
the MHC HaploTYPE Consortium and collaborators. Further information  
can be found at

http://www.sanger.ac.uk/HGP/Chr6/MHC

DASS-94G7 is from a DNA-arts SSTO human bac library VECTOR:

pBelOBAC11.

## FEATURES

## source

```

1..19937
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/clone="DASS-94G7"
/clone_lib="DNA-arts-BAC.1-SSTO.1"

```

## ORIGIN

```

Query Match      98.1%; Score 83.4; DB 8; Length 19937;
Best Local Similarity 98.8%; Pred. No. 5.6e-18;
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 60
    |||||
Db 19544 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 19485
    |||||

QY 61 CTGGGTGCCCCACTAAGTAGAATAA 85
    |||||
Db 19484 CTGGGTGCCCCACTAAGTAGAATAA 19460
    |||||

```

## RESULT 6

```

CT009502/c
LOCUS
DEFINITION
Human DNA sequence from clone DAMC-2718 on chromosome 6, complete
sequence.
ACCESSION
CT009502
VERSION
CT009502.2 GI:68644850
KEYWORDS
HTG.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

```

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

COMMENT

Hummeridae; Homo.  
1 (bases 1 to 35167)  
Palmer, S.  
Direct Submission  
Submitted (30-JUN-2005) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk  
Clone requests: clonerequest@sanger.ac.uk  
On Jul 5, 2005 this sequence version replaced gi:68226286.  
-----  
Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: http://www.sanger.ac.uk  
Contact: vegas@sanger.ac.uk  
-----  
This sequence was finished as follows unless otherwise noted: all  
regions were either double-stranded or sequenced with an alternate  
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by at least  
one subclone; and the assembly was confirmed by restriction digest,  
except on the rare occasion of the clone being a YAC.  
The following abbreviations are used to associate primary accession  
numbers given in the feature table with their source databases:  
Em., EMBL; Sw., SWISSPROT; Tr., TrEMBL; Wp., WORMPEP; Information  
on the WORMPEP database can be found at  
http://www.sanger.ac.uk/Projects/C\_elegans/wormpep This sequence  
was generated from part of bacterial clone contigs constructed by  
the MHC Haplotype Consortium and collaborators. Further information  
can be found at  
http://www.sanger.ac.uk/HGP/Chr6/MHC  
DAMC-2718 is from the DNA-Arts.org BAC library MCF.1 VECTOR:  
pBelOBAC11.  
-----  
FEATURES  
source  
Location/Qualifiers  
1..35167  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="6"  
/clone="DAMC-2718"  
/clone\_lib="DNA-Arts.org BAC library MCF.1"  
-----  
ORIGIN  
Query Match 98.1%; Score 83.4; DB 8; Length 35167;  
Best Local Similarity 98.8%; Pred. No. 5.6e-18;  
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAGCATTTCCCGGAGGAGTTT 60  
Db 31062 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAGCATTTCCCGGAGGAGTTT 31003  
Qy 61 CTGGTGCCCGCCACTAAGTAGAAAA 85  
Db 31002 CTGGTGCCCGCCACTAAGTAGAATAA 30978  
-----  
RESULT 7  
CR753889/C  
LOCUS  
DEFINITION  
Human DNA sequence from clone DADB-78A10 on chromosome 6, complete  
sequence.  
ACCESSION  
CR753889  
VERSION  
CR753889.4 GI:52421005  
KEYWORDS  
HTG.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
1 (bases 1 to 47546)  
Brown, J.  
Direct Submission  
Submitted (21-SEP-2004) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:

humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk  
On Sep 21, 2004 this sequence version replaced gi:51988278.  
-----  
Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: http://www.sanger.ac.uk  
Contact: humquery@sanger.ac.uk  
-----  
During sequence assembly data is compared from overlapping clones.  
Where differences are found these are annotated as variations  
together with a note of the overlapping clone name. Note that the  
variation annotation may not be found in the sequence submission  
corresponding to the overlapping clone, as we submit sequences with  
only a small overlap as described above.  
This sequence was finished as follows unless otherwise noted: all  
regions were either double-stranded or sequenced with an alternate  
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by at least  
one plasmid subclone or more than one M13 subclone; and the  
assembly was confirmed by restriction digest, except on the rare  
occasion of the clone being a YAC.  
The following abbreviations are used to associate primary accession  
numbers given in the feature table with their source databases:  
Em., EMBL; Sw., SWISSPROT; Tr., TrEMBL; Wp., WORMPEP; Information  
on the WORMPEP database can be found at  
http://www.sanger.ac.uk/Projects/C\_elegans/wormpep This sequence  
was generated from part of bacterial clone contigs of human  
chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping  
Group. Further information can be found at  
http://www.sanger.ac.uk/HGP/Chr6  
DADB-78A10 is from a DNA-arts DBB human bac library VECTOR:  
pBelOBAC11.  
-----  
FEATURES  
source  
Location/Qualifiers  
1..47546  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="6"  
/clone="DADB-78A10"  
/clone\_lib="DNA-arts-BAC.1-DBB.1"  
-----  
ORIGIN  
Query Match 98.1%; Score 83.4; DB 8; Length 47546;  
Best Local Similarity 98.8%; Pred. No. 5.6e-18;  
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAGCATTTCCCGGAGGAGTTT 60  
Db 24535 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAGCATTTCCCGGAGGAGTTT 24476  
Qy 61 CTGGTGCCCGCCACTAAGTAGAAAA 85  
Db 24475 CTGGTGCCCGCCACTAAGTAGAATAA 24451  
-----  
RESULT 8  
AL935043/C  
LOCUS  
DEFINITION  
Human DNA sequence from clone DAQB-69D7 on chromosome 6 Contains  
the TAP2 gene for transporter 2, ATP-binding cassette, sub-family B  
(MDR/TAP), the PSMB8 gene for proteasome (prosome, macropain)  
subunit, beta type, 8 (large multifunctional protease 7), the TAP1  
gene for transporter 1, ATP-binding cassette, sub-family B  
(MDR/TAP), the PSMB9 gene for proteasome (prosome, macropain)  
subunit, beta type, 9 (large multifunctional protease 2), the  
PPP1R2p1 gene for protein phosphatase 2, regulatory (inhibitor)  
subunit 2 pseudogene 1 and 1 CpG island, complete sequence.  
ACCESSION  
AL935043  
VERSION  
AL935043.4 GI:24430271  
KEYWORDS  
HTG; CpG island; PPP1R2p1; PSMB8; PSMB9; TAP1; TAP2.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 63179)  
Tracey,A.  
Direct Submission  
Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk  
Clone requests: clonerequest@sanger.ac.uk  
On Oct 30, 2002 this sequence version replaced gi:24415778.  
The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:  
Emi, EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at  
http://www.sanger.ac.uk/Projects/C\_elegans/wormpep  
This sequence was generated from part of bacterial clone contigs constructed by the MHC Haplotype Consortium and collaborators. Further information can be found at  
http://www.sanger.ac.uk/HGP/Chr6/MHC  
----- Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: http://www.sanger.ac.uk  
Contact: vegas@sanger.ac.uk  
-----

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.  
DAQB-69D7 is from a DNA-arts QBL human bac library VECTOR: pBelobAC11.

FEATURES  
source

Location/Qualifiers  
1..63179  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="6"  
/clone="DAQB-69D7"  
/clone\_lib="DNA-arts-BAC.1-QBL.1"

misc\_feature

1  
/note="Clone left end: DAQB-69D7"  
complement(join(1711..2196,9277..9413,9807..9966,10144..10317,10495..10683,12210..12338,12504..12701,15031..15236,15520..15650,17414..17528,17618..18114,18532..18623))  
/gene="TAP2"

mRNA

/locus\_tag="DAQB-69D7.2-002"  
complement(join(1711..2196,9277..9413,9807..9966,10144..10317,10495..10683,12210..12338,12504..12701,15031..15236,15520..15650,17414..17528,17618..18114,18532..18623))  
/gene="TAP2"  
/locus\_tag="DAQB-69D7.2-002"  
/product="transporter 2, ATP-binding cassette, sub-family B (MDR)/TAP"  
/note="match: cDNAs: AF078671.1 AF105151.1"  
complement(1723..1728)

polyA\_signal

/locus\_tag="DAQB-69D7.2-002"  
complement(join(2167..2196,9277..9413,9807..9966,10144..10317,10495..10683,12210..12338,12504..12701,15031..15236,15520..15650,17414..17528,17618..18110))  
/gene="TAP2"

CDS

/locus\_tag="DAQB-69D7.2-002"  
complement(join(2167..2196,9277..9413,9807..9966,10144..10317,10495..10683,12210..12338,12504..12701,15031..15236,15520..15650,17414..17528,17618..18110))  
/gene="TAP2"  
/locus\_tag="DAQB-69D7.2-002"  
/standard\_name="OTTHUMP0000014665"  
/note="match: cDNAs: O95410 P36372 Q03519 Q28438"  
/codon\_start=1  
/product="transporter 2, ATP-binding cassette, sub-family B (MDR)/TAP"  
/protein\_id="CAI18621.1"

/db\_xref="GI:56207306"  
/db\_xref="GGA:Q9UP03"  
/db\_xref="InterPro:IPR001140"  
/db\_xref="InterPro:IPR003439"  
/db\_xref="InterPro:IPR003593"  
/db\_xref="InterPro:IPR005293"  
/db\_xref="InterPro:IPR011527"  
/db\_xref="UniProt/TrEMBL:Q9UP03"  
/translation="MRLPDLRPMTSLLLVDAALLWLLQGLPLGLTLLPQGLPGLMEGLTL  
LVGYGAAGLSWSLWAVLSPGQAEQKQDQNNKVMRLIKLSRPDLPLVAAPFFVL  
LAVIGETLIPIHSGRVIDILGGDFDPHAFASAIFFMCLFSGSLGACGCGCTYTM  
SRINLRIQGLFSSLLRQDLGFFQETKTGELNSRLSDTTLMSNMLPLNANVLLRSIV  
KVGLYGMFLSIPRLTLLSLHMPPTIAAEKYNTNRHQBVLREIQAIVARAGQVRE  
AVGGLQTVSPGAEHEVCYKEALQCQQLYWRDLERALLYLVRVHLGVQMLML  
SCGLQMQGELUTQGSLTSMFYQESVSVQTLVIYIGDMLSNVGAEEKVFYMQRQ  
PNLSPGTLAPTTLQGVVRFQDVSPAYPNRPDPVLKGLTFTLRPEVETALVPGNSGG  
KSTVAALLQNYOPTGGVLLDEKPIQSQEHCVLHQSQVSVGQEPVLFSGSVRNNTAY  
GLQSCDDKVMAAAQAHAADDFIQEMEHGIYTDVGEKSQLAAGQKQRAIAIALVRD  
PRVLLDEATSLDVQCEQAQKTLKFMIP"  
complement(join(8366..8911,9277..9413,9807..9966,10144..10317,10495..10683,12210..12338,12504..12701,15031..15236,15520..15650,17414..17528,17618..18114,18532..18623))  
/gene="TAP2"  
/locus\_tag="DAQB-69D7.2-001"  
complement(join(8366..8911,9277..9413,9807..9966,10144..10317,10495..10683,12210..12338,12504..12701,15031..15236,15520..15650,17414..17528,17618..18114,18532..18623))  
/gene="TAP2"  
/locus\_tag="DAQB-69D7.2-001"  
/product="transporter 2, ATP-binding cassette, sub-family B (MDR)/TAP"  
/note="match: cDNAs: BC002751.1 M74447.1 M84748.1 222936.1"  
complement(8366)  
/gene="TAP2"  
/locus\_tag="DAQB-69D7.2-001"  
complement(join(8783..8911,9277..9413,9807..9966,10144..10317,10495..10683,12210..12338,12504..12701,15031..15236,15520..15650,17414..17528,17618..18110))  
/gene="TAP2"  
/locus\_tag="DAQB-69D7.2-001"  
/standard\_name="OTTHUMP0000014664"  
/note="match: proteins: O95410 P36372 Q03519 Q28438"  
/codon\_start=1  
/product="transporter 2, ATP-binding cassette, sub-family B (MDR)/TAP"  
/protein\_id="CAI18622.1"  
/db\_xref="GI:56207307"  
/db\_xref="InterPro:IPR001140"  
/db\_xref="InterPro:IPR003439"  
/db\_xref="InterPro:IPR003593"  
/db\_xref="InterPro:IPR005293"  
/translation="MRLPDLRPMTSLLLVDAALLWLLQGLPLGLTLLPQGLPGLMEGLTL  
LVGYGAAGLSWSLWAVLSPGQAEQKQDQNNKVMRLIKLSRPDLPLVAAPFFVL  
LAVIGETLIPIHSGRVIDILGGDFDPHAFASAIFFMCLFSGSLGACGCGCTYTM  
SRINLRIQGLFSSLLRQDLGFFQETKTGELNSRLSDTTLMSNMLPLNANVLLRSIV  
KVGLYGMFLSIPRLTLLSLHMPPTIAAEKYNTNRHQBVLREIQAIVARAGQVRE  
AVGGLQTVSPGAEHEVCYKEALQCQQLYWRDLERALLYLVRVHLGVQMLML  
SCGLQMQGELUTQGSLTSMFYQESVSVQTLVIYIGDMLSNVGAEEKVFYMQRQ  
PNLSPGTLAPTTLQGVVRFQDVSPAYPNRPDPVLKGLTFTLRPEVETALVPGNSGG  
KSTVAALLQNYOPTGGVLLDEKPIQSQEHCVLHQSQVSVGQEPVLFSGSVRNNTAY  
GLQSCDDKVMAAAQAHAADDFIQEMEHGIYTDVGEKSQLAAGQKQRAIAIALVRD  
PRVLLDEATSLDVQCEQAQKTLKFMIP"  
complement(join(20595..20925,21409..21613,22012..22141,22550..22661,22820..22967,24157..24581))  
/gene="P5MB8"  
/locus\_tag="DAQB-69D7.6-001"  
complement(join(20595..20925,21409..21613,22012..22141,

```
22550. .22661,22820. .22967,24157. .24581))
/gene="PSMB8"
/locus tag="DAQB-69D7.6-001"
/product="proteasome (prosome, macropain) subunit, beta
type, 8 (large multifunctional protease 7)"
/notes="match: ESTs: BM547457.1 BU935203.1
match: cDNAs: BC001114.1 X62598.1"
complement(join(20595. .20925,21409. .21613,22012. .22141,
22550. .22661,22820. .22967,23728. .23925))
/gene="PSMB8"
/locus tag="DAQB-69D7.6-002"
complement(join(20595. .20925,21409. .21613,22012. .22141,
22550. .22661,22820. .22967,23728. .23925))
/gene="PSMB8"
/locus tag="DAQB-69D7.6-002"
/product="proteasome (prosome, macropain) subunit, beta
type, 8 (large multifunctional protease 7)"
/notes="match: ESTs: BM547457.1
match: cDNAs: AK031801.1 U17496.1 U17497.1 U22033.1"
complement(20595)
/gene="PSMB8"
/locus tag="DAQB-69D7.6-002"
complement(20613. .20618)
/gene="PSMB8"
/locus tag="DAQB-69D7.6-002"
complement(join(20776. .20925,21409. .21613,22012. .22141,
22550. .22589,22820. .22967,23728. .23901))
/gene="PSMB8"
/locus tag="DAQB-69D7.6-003"
complement(join(20776. .20925,21409. .21613,22012. .22141,
22550. .22589,22820. .22967,23728. .23901))
/gene="PSMB8"
/locus tag="DAQB-69D7.6-003"
/product="proteasome (prosome, macropain) subunit, beta
type, 8 (large multifunctional protease 7)"
/notes="match: ESTs: BQ685477.1 BQ686592.1"
complement(join(20837. .20925,21409. .21613,22012. .22141,
22550. .22661,22820. .22967,24157. .24291))
/gene="PSMB8"

Query Match 98.1%; Score 83.4; DB 8; Length 63179;
Best Local Similarity 98.8%; Pred. No. 5.6e-18;
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGTGAAGTGCACTCTCTGGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 60
Db 20698 CGGTGAAGTGCACTCTCTGGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 20639

Qy 61 CTGGGTGCCCCCACTAAGTAGAAAA 85
Db 20638 CTGGGTGCCCCCACTAAGTAGAAAA 20614

RESULT 9
AX330551 66109 bp DNA linear PAT 09-JAN-2002
LOCUS
DEFINITION Sequence 1060 from Patent WO0194629.
ACCESSION AX330551
VERSION AX330551.1 GI:19103529
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
AUTHORS Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
Horvath, S., Soppet, D.R. and Weaver, Z.
TITLE Cancer gene determination and therapeutic screening using signature
gene sets
JOURNAL Patent: WO 0194629-A 1060 13-DEC-2001;
Avalon Pharmaceuticals (US)
FEATURES
Location/Qualifiers
```

```
source 1. .66109
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 98.1%; Score 83.4; DB 6; Length 66109;
Best Local Similarity 98.8%; Pred. No. 5.6e-18;
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGTGAAGTGCACTCTCTGGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 60
Db 37840 CGGTGAAGTGCACTCTCTGGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 37899

Qy 61 CTGGGTGCCCCCACTAAGTAGAAAA 85
Db 37900 CTGGGTGCCCCCACTAAGTAGAAAA 37924

RESULT 10
AX411124 66109 bp DNA linear PAT 14-JUN-2002
LOCUS
DEFINITION Sequence 3771 from Patent WO0229103.
ACCESSION AX411124
VERSION AX411124.1 GI:21443829
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
AUTHORS Alvares, C., Horne, D., Peres-da-Silva, S. and Vockley, J.G.
TITLE Gene expression profiles in liver cancer
JOURNAL Patent: WO 0229103-A 3771 11-APR-2002;
GENE LOGIC INC (US)
FEATURES
Location/Qualifiers
source 1. .66109
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/notes="EMBL/GenBank Accession No. X66401"

ORIGIN

Query Match 98.1%; Score 83.4; DB 6; Length 66109;
Best Local Similarity 98.8%; Pred. No. 5.6e-18;
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGTGAAGTGCACTCTCTGGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 60
Db 37840 CGGTGAAGTGCACTCTCTGGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 37899

Qy 61 CTGGGTGCCCCCACTAAGTAGAAAA 85
Db 37900 CTGGGTGCCCCCACTAAGTAGAAAA 37924

RESULT 11
HSMHCAPG 66109 bp DNA linear PRI 18-APR-2005
LOCUS
DEFINITION H.sapiens genes TAP1, TAP2, LMP2, LMP7 and DOB.
ACCESSION X66401 S57528
VERSION X66401.1 GI:34634
KEYWORDS DOB gene; LMP2 gene; LMP7 gene; major histocompatibility complex
class II; tap1 gene; tap2 gene.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
AUTHORS Beck, S., Kelly, A., Radley, E., Khurshid, F., Alderton, R.P. and
Trowsdale, J.
```

**TITLE** DNA sequence analysis of 66 kb of the human MHC class II region encoding a cluster of genes for antigen processing

**JOURNAL** J. Mol. Biol. 228 (2), 433-441 (1992)

**PUBMED** 1453454

**REFERENCE** 2 (bases 1 to 66109)

**AUTHORS** Glynn, R., Kerr, L.A., Mockridge, I., Beck, S., Kelly, A. and Towdale, J.

**TITLE** The major histocompatibility complex-encoded proteasome component LMP7: alternative first exons and post-translational processing

**JOURNAL** Eur. J. Immunol. 23 (4), 860-866 (1993)

**PUBMED** 8458375

**REFERENCE** 3 (bases 1 to 66109)

**AUTHORS** Beck, S.

**TITLE** Direct Submission

**JOURNAL** Submitted (21-MAY-1992) S. Beck, Imperial Cancer Research Fund, ICRF, 44 Lincoln's Inn Fields, London WC2A 3PX, UK

**COMMENT** On Sep 23, 2004 this sequence version replaced gi:301893.

**FEATURES** Location/Qualifiers

**source** 1..66109

    /organism="Homo sapiens"

    /mol\_type="genomic DNA"

    /db\_xref="taxon:9606"

    /chromosome="6"

    <1..>66109

    /notes="classII-J-box"

    classII-X-box

    classII-Y-box

        1..272

            /rpt\_family="ALU"

            2517..2811

            /rpt\_family="ALU"

            2927..4135

            /rpt\_family="LINE"

            4136..4864

            /rpt\_family="LINE"

            4662..4864

            /notes="mer13"

            /rpt\_family="MER"

            5182..5514

            /rpt\_family="ALU"

            5525..5876

            /rpt\_family="LINE"

            5672..5860

            /notes="mer13"

            /rpt\_family="MER"

            6391..6689

            /rpt\_family="ALU"

            8843..9091

            /rpt\_family="ALU"

            13973..14135

            /rpt\_family="ALU"

            14542..14795

            /rpt\_family="ALU"

            14820..15595

            /rpt\_family="LINE"

            15650..15789

            /rpt\_family="ALU"

            15907..16194

            /rpt\_family="ALU"

            16853..17150

            /rpt\_family="ALU"

            18240..18529

            /rpt\_family="ALU"

            complement(join(19127..19254,20154..20295,20525..20654,21265..21396,22455..22522,24371..24430))

            /genes="LMP2"

            complement(join(19127..19254,20154..20295,20525..20654,21265..21396,22455..22522,24371..24430))

            /genes="LMP2"

            /functions="antigen processing"

            /notes="proteasome related"

            /codon\_start=1

            /protein\_id="CAA47024.1"

    /db\_xref="GI:34635"

    /db\_xref="GOA:P28065"

    /db\_xref="InterPro:IPR000243"

    /db\_xref="InterPro:IPR001353"

    /db\_xref="UniProt/Swiss-Prot:P28065"

    translation="MLRAGAPTGLPRAGEVHTTITMAVEFDGGVVMGSDSRVSAGEAVNRRVFDKLSPLHERIYCALSGSADADAQADMAAYQLELHGIELEEEPLVLAANVVRNISKYREDLSAHLNVAQMDQREGQVGTGLGMLTRQPPFAIGSGSFTIYGYDAAYKPGMSPECCRRFTTDAIALAMSRDSSGGVIYLVTTAAGVDHRVILGNELPKFYD E"

        complement(19127..19254)

        /genes="LMP2"

        /number=6

        complement(19253..20155)

        /genes="LMP2"

        /number=5

        19588..19910

        /rpt\_family="ALU"

        complement(20154..20295)

        /genes="LMP2"

        /number=5

        complement(20294..20524)

        /genes="LMP2"

        /number=4

        complement(20525..20654)

        /genes="LMP2"

        /number=4

        complement(20653..21266)

        /genes="LMP2"

        /number=3

        complement(21265..21396)

        /genes="LMP2"

        /number=3

        complement(21395..22456)

        /genes="LMP2"

        /number=2

        21438..21741

        /rpt\_family="ALU"

        complement(22455..22522)

        /genes="LMP2"

        /number=2

        complement(22521..24372)

        /genes="LMP2"

        /number=1

        complement(24011..24707)

        complement(24371..24430)

        /genes="LMP2"

        /number=1

        24506..24519

        misc\_signal

        24826..24838

        /notes="ISRE"

        24896..24909

        25021..25025

        join(25024..25621,26158..26272,26421..26551,27511..27716,28143..28340,29542..29670,29820..30008,30568..30741,30985..31147,31456..31592,32875..33081)

        /genes="TAP1"

        join(25024..25621,26158..26272,26421..26551,27511..27716,28143..28340,29542..29670,29820..30008,30568..30741,30985..31147,31456..31592,32875..33081)

        /genes="TAP1"

        /function="antigen processing"

        /notes="peptide pump"

        /codon\_start=1

        /product="ABC-transporter"

        /protein\_id="CAA47025.1"

        /db\_xref="GI:34636"

        /db\_xref="GOA:Q03518"

        /db\_xref="InterPro:IPR001140"

        /db\_xref="InterPro:IPR003439"

        /db\_xref="InterPro:IPR003593"

        /db\_xref="InterPro:IPR005293"

        /db\_xref="InterPro:IPR009078"

        CDS

            join(25024..25621,26158..26272,26421..26551,27511..27716,28143..28340,29542..29670,29820..30008,30568..30741,30985..31147,31456..31592,32875..33081)

            /genes="TAP1"

            /function="antigen processing"

            /notes="peptide pump"

            /codon\_start=1

            /product="ABC-transporter"

            /protein\_id="CAA47025.1"

            /db\_xref="GI:34636"

            /db\_xref="GOA:Q03518"

            /db\_xref="InterPro:IPR001140"

            /db\_xref="InterPro:IPR003439"

            /db\_xref="InterPro:IPR003593"

            /db\_xref="InterPro:IPR005293"

            /db\_xref="InterPro:IPR009078"

```
/db_xref="UniProt/Swiss-Prot:Q03518"
/translation="MASSRCAPRGRCCLFASLAWLGTVLLLLADWVLLRTALPRIF
SLIVPTALPLLRWAVGLSRWAVLWLGACGLVLRATVGSKEAGACQWLAALKPLAAA
LGLALPGALFRELISWAGPSADSTRLLHWSHPTAFVVSAAALPAALWHLGLSL
WFGGQGGSNPVRLLGLGSLTRLSLEVLVLSLGGEMAIPTFTGRLTDWILQD
GSADFTFRNLTLMSILITIASAVLEFVGDGIYNNMGHVSHLQGEVGVLRQETEFF
QQQNTGIMSRSVDTSTLSDSLSENLSLFLVLYRGLCLLGLMLWGSVSLTVTLIT
LPLFLPKKQWYQLLEVQVRESLAKSQVAIEALSAMPTVRSFANESEGAQKPRE
KLOIKTLNOKRAVAVNSWTTISGMLKVLGILYTGQLVTSQAVSSGNLVTFVLY
QMOFTQAVEVLLSTYPRQVAGSSEKIPEYLDTRPCRPGLLTPHLHGLVQFQDV
SFAYPNRPDLVLTQGLFTLRPGEVTLVGPNGSGKSTVAALLQNLVQPTGQGLLDG
KPLPQYHRILHQAVAGQEPQVFGSLQENIAYGLTQKPTWEEITAAAVKGAHSP
ISGLPQGYDTEVDEAGSQLSGQQRQAVALARIRKPCVLILDDATSALDANSOLOVE
QLLYSPERYSRVLLITQHLSLVEQADHILFLEGGAIREGGTHQQLMEKKGCYWAV
QAPADAPE"
25024..25621
/gene="TAP1"
/number=1
25622..26157
/gene="TAP1"
/number=1
26158..26272
/gene="TAP1"
/number=2
26273..26420
/gene="TAP1"
/number=2
26421..26551
/gene="TAP1"
/number=3
26552..27510
/gene="TAP1"
/number=3
26643..26752
/number=3
26643..26752
/number="ca"
/rpt_unit="26643..26644"
26919..26967
/gene="TAP1"
/rpt_family="LAD"
27004..27064
repeat_region
26919..26967
repeat_unit
27004..27064
repeat_unit
```

```
Query Match 98.1%; Score 83.4; DB 8; Length 66109;
Best Local Similarity 98.8%; Pred. No. 5.6e-18;
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGTGAAGTGCATCTTCGCGTGTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 60
Db 37840 CGGTGAAGTGCATCTTCGCGTGTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 37899

Qy 61 CTGGTGCCCCCACTAAGTAGAAA 85
Db 37900 CTGGTGCCCCCACTAAGTAGAATAA 37924
```

```
RESULT 12
CR762476/c CR762476 95554 bp DNA linear PRI 19-OCT-2004
LOCUS Human DNA sequence from clone DAAP-57C1 on chromosome 6, complete
DEFINITION sequence.
ACCESSION CR762476
VERSION CR762476.2 GI:54114714
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 95554)
AUTHORS Wood, J.
TITLE Direct Submission
JOURNAL Submitted (19-OCT-2004) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
```

## COMMENT

On Oct 13, 2004 this sequence version replaced gi:52000322.  
----- Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: <http://www.sanger.ac.uk>  
Contact: [humquery@sanger.ac.uk](mailto:humquery@sanger.ac.uk)  
-----

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Sw., SWISSPROT; Tr., TrEMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at

[http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) This sequence was generated from part of bacterial clone contigs constructed by the MHC Haplotype Consortium and collaborators. Further information can be found at

<http://www.sanger.ac.uk/HGP/Chr6/MHC>  
DAAP-57C1 is from a APD human bac - B Lymphoblastoid Cell Line  
Library VECTOR: pBelobAC11.

## FEATURES

```
source
1..95554
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/clone="DAAP-57C1"
/clone_lib="APD"
```

## ORIGIN

```
Query Match 98.1%; Score 83.4; DB 8; Length 95554;
Best Local Similarity 98.8%; Pred. No. 5.5e-18;
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 1 CGGTGAAGTGCATCTTCGCGTGTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 60
Db 68596 CGGTGAAGTGCATCTTCGCGTGTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 68537
```

```
Qy 61 CTGGTGCCCCCACTAAGTAGAAA 85
Db 68536 CTGGTGCCCCCACTAAGTAGAATAA 68512
```

## RESULT 13

```
BX927138/c BX927138 110856 bp DNA linear PRI 10-MAR-2004
LOCUS Human DNA sequence from clone DAMA-91123 on chromosome 6, complete
DEFINITION sequence.
```

ACCESSION BX927138

VERSION BX927138.5 GI:45379240

KEYWORDS HTG.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE 1 (bases 1 to 110856)

AUTHORS Sycamore, N.

TITLE Direct Submission

JOURNAL Submitted (09-MAR-2004) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
[humquery@sanger.ac.uk](mailto:humquery@sanger.ac.uk)



COMMENT  
 humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk  
 On Mar 11, 2004 this sequence version replaced gi:42820837.  
 ----- Genome Center  
 Center: Wellcome Trust Sanger Institute  
 Web site: http://www.sanger.ac.uk  
 Contact: humquery@sanger.ac.uk  
 -----  
 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.  
 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.  
 The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C\_elegans/wormpep This sequence was generated from part of bacterial clone contigs constructed by the MHC Haplotype Consortium and collaborators. Further information can be found at  
 http://www.sanger.ac.uk/HGP/Chr6/MHC  
 DANA-91123 is from the DNA-Arts human BAC library MANN.1 VECTOR: pBelOBAC11.  
 ----- Location/Qualifiers  
 source  
 1. .110856  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="6"  
 /clone="DANA-91123"  
 /clone\_lib="DNA-Arts BAC library MANN.1"  
 -----  
 ORIGIN  
 Query Match 98.1%; Score 83.4; DB 8; Length 110856;  
 Best Local Similarity 98.8%; Pred. No. 5.5e-18;  
 Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCCAGGGAAGTTT 60  
 |||||  
 Db 8030 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCCAGGGAAGTTT 7971  
 |||||  
 QY 61 CTGGTGCCCCCACTAAGTAGAAAA 85  
 |||||  
 Db 7970 CTGGTGCCCCCACTAAGTAGAATAA 7946  
 |||||  
 RESULT 14  
 BX088556/c  
 LOCUS BX088556 114576 bp DNA linear PRI 18-MAY-2005  
 DEFINITION Human DNA sequence from clone DASS-67H14 on chromosome 6, complete sequence.  
 ACCESSION BX088556  
 VERSION BX088556.11 GI:32968209  
 KEYWORDS HTG;  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.  
 1 (bases 1 to 114576)  
 Sycamore.N.  
 Direct Submission  
 Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton, JOURNAL

Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk  
 Clone requests: clonerequest@sanger.ac.uk  
 On Jul 18, 2003 this sequence version replaced gi:32563203.  
 The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C\_elegans/wormpep This sequence was generated from part of bacterial clone contigs constructed by the MHC Haplotype Consortium and collaborators. Further information can be found at  
 http://www.sanger.ac.uk/HGP/Chr6/MHC  
 DASS-67H14 is from a DNA-Arts SSTS human bac library VECTOR: pBelOBAC11  
 ----- Genome Center  
 Center: Wellcome Trust Sanger Institute  
 Center code: SC  
 Web site: http://www.sanger.ac.uk  
 Contact: vegas@sanger.ac.uk  
 -----  
 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.  
 ----- Location/Qualifiers  
 source  
 1. .114576  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="6"  
 /clone="DASS-67H14"  
 /clone\_lib="DNA-Arts-BAC.1-SSTS.1"  
 -----  
 misc\_feature  
 1  
 /notes="Clone\_left\_end: DASS-67H14"  
 misc\_feature  
 114576  
 /notes="Clone\_right\_end: DASS-67H14"  
 -----  
 ORIGIN  
 Query Match 98.1%; Score 83.4; DB 8; Length 114576;  
 Best Local Similarity 98.8%; Pred. No. 5.5e-18;  
 Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCCAGGGAAGTTT 60  
 |||||  
 Db 1607 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCCAGGGAAGTTT 1548  
 |||||  
 QY 61 CTGGTGCCCCCACTAAGTAGAAAA 85  
 |||||  
 Db 1547 CTGGTGCCCCCACTAAGTAGAATAA 1523  
 |||||  
 RESULT 15  
 CR786563/c  
 LOCUS CR786563 116859 bp DNA linear HTG 26-OCT-2004  
 DEFINITION Homo sapiens chromosome 6 clone DADB-30110, 2 unordered pieces.  
 ACCESSION CR786563  
 VERSION CR786563.5 GI:54633101  
 KEYWORDS HTG; HTGS PHASE1; HTGS\_CANCELLED.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.  
 1 (bases 1 to 116859)  
 Sims.S.  
 Direct Submission  
 Submitted (22-OCT-2004) Wellcome Trust Sanger Institute, Hinxton, JOURNAL  
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk  
 On Oct 25, 2004 this sequence version replaced gi:54306012.



```
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: b0B30110
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 78856 bases at least Q40
Consensus quality: 78857 bases at least Q30
Consensus quality: 78857 bases at least Q20
Insert size: 116759; sum-of-contigs
Insert size: 84033; 2.3% error; agarose-fp
Quality coverage: 6.62x in Q20 bases; sum-of-contigs Quality
coverage: 9.20x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 78857: contig of 78857 bp in length
*
* 78858 78957: gap of 100 bp
*
* 78958 116859: contig of 37902 bp in length.
FEATURES
source
    1..116859
        /organism="Homo sapiens"
        /mol_type="genomic DNA"
        /db_xref="taxon:9606"
        /chromosome="6"
        /clone="DADB-30110"
        /clone_lib="DNA-arts-BAC.1-DBB.1"
    1..78857
        /note="assembly_fragment:01987"
    78958..116859
        /note="assembly_fragment:02008"
misc_feature
misc_feature
ORIGIN
Query Match          98.1%; Score 83.4; DB 14; Length 116859;
Best Local Similarity 98.8%; Pred. No. 5,5e-18;
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy      1  CGGTGAAGTGCATCTTCTGCGTGTCTCTATTGAAACAAGCATTTCCTCCCGAGGAAAGTTT 60
Db      105184 CGGTGAAGTGCATCTTCTGCGTGTCTCTATTGAAACAAGCATTTCCTCCCGAGGAAAGTTT 105125
Qy      61  CTGGGTGCCCTACTAGTAGAATAA 85
Db      105124 CTGGGTGCCCTACTAGTAGAATAA 105100
Search completed: December 7, 2005, 04:22:18
Job time : 753.729 secs
```

***This Page Blank (ucp)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 11:31:40 ; Search time 211.652 Seconds  
(without alignments)  
3320.997 Million cell updates/sec

Title: US-09-980-046B-10  
Perfect score: 85  
Sequence: 1 cggtagagtcattcttcg.....tgccccactaagtagagaaaa 85

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA\_Main:  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:  
2: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:  
3: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:  
4: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:  
5: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:  
6: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:  
7: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:  
8: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:  
9: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:  
10: /cgn2\_6/ptodata/1/pubpna/US11\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	83.4	98.1	1126	8	US-10-741-600-161
2	83.4	98.1	1143	3	US-09-925-302-79
3	83.4	98.1	1143	3	US-09-925-302-79
4	83.4	98.1	1271	5	US-10-084-817-272
5	83.4	98.1	1320	6	US-10-159-563-162
6	83.4	98.1	1539	8	US-10-741-600-162
7	83.4	98.1	2027	5	US-10-198-846-10394
8	83.4	98.1	3674	5	US-10-084-817-258
9	83.4	98.1	16175	8	US-10-741-600-17596
10	83.4	98.1	20773	8	US-10-741-600-17595
11	83.4	98.1	28789	8	US-10-741-600-17894
12	83.4	98.1	66109	3	US-09-880-107-3768
13	83.4	98.1	66109	3	US-09-873-367C-1060
14	83.4	98.1	66109	9	US-10-843-641A-1060
15	83.4	98.1	198285	8	US-09-880-107-3814
16	83.4	98.1	198285	8	US-10-775-169-338
17	83.4	98.1	198285	9	US-10-756-149-12
18	54.6	64.2	11398	6	US-10-311-455-2411
19	54.6	64.2	11398	6	US-10-311-455-2412
20	54.6	64.2	3673778	6	US-10-312-841-1
21	54.6	64.2	3673778	6	US-10-312-841-2
22	28.2	33.2	1530	4	US-09-925-065A-52319
23	28.2	33.2	1530	4	US-09-925-065A-52320

ALIGNMENTS

RESULT 1

US-10-741-600-161  
; Sequence 161, Application US/10741600  
; Publication No. US20050026169A1  
; GENERAL INFORMATION:  
; APPLICANT: CAROULLI, Michele et al.  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001499  
; CURRENT APPLICATION NUMBER: US/10741.600  
; CURRENT FILING DATE: 2003-12-22  
; NUMBER OF SEQ ID NOS: 73997  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 161  
; LENGTH: 1126  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-741-600-161

Query Match 98.1%; Score 83.4; DB 8; Length 1126;  
Best Local Similarity 98.8%; Pred. No. 3.8e-20;  
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGTGAAGTGCATCTTCTGCGTGTCTCTATTTCGAACAAGCATTTCCCCAGGAAGTTT 60  
Db 1022 CGGTGAAGTGCATCTTCTGCGTGTCTCTATTTCGAACAAGCATTTCCCCAGGAAGTTT 1081  
Qy 61 CTGGGTGCCCCCACTAAGTAGAGAAAAA 85  
Db 1082 CTGGGTGCCCCCACTAAGTAGAGATAA 1106

RESULT 2

US-09-925-302-79  
; Sequence 79, Application US/0925302  
; Patent No. US2002004941A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA104  
; CURRENT APPLICATION NUMBER: US/09/925.302  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05918  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12

```
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 79
; LENGTH: 1143
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1126)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-302-79

Query Match      98.1%; Score 83.4; DB 3; Length 1143;
Best Local Similarity 98.8%; Pred. No. 3.8e-20;
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 60
DB 1006 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 1065

QY 61 CTGGTGCCCCCACTAAGTAGAATAA 85
DB 1066 CTGGTGCCCCCACTAAGTAGAATAA 1090

RESULT 3
US-09-925-302-79
Sequence 79, Application US/09925302
Publication No. US20030064072A9
GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 79
; LENGTH: 1143
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1126)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-302-79

Query Match      98.1%; Score 83.4; DB 3; Length 1143;
Best Local Similarity 98.8%; Pred. No. 3.8e-20;
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 60
DB 1006 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 1065

QY 61 CTGGTGCCCCCACTAAGTAGAATAA 85
DB 1066 CTGGTGCCCCCACTAAGTAGAATAA 1090

RESULT 4
US-10-084-817-272
Sequence 272, Application US/10084817
Publication No. US20030119009A1
GENERAL INFORMATION:
; APPLICANT: Susan Stuart
; APPLICANT: Jed G. Nuchtern
; APPLICANT: Sharon E. Plon
; APPLICANT: Jason M. Shohet
```

```
; TITLE OF INVENTION: GENES REGULATED BY MYCN ACTIVATION
; FILE REFERENCE: PA-0046 US
; CURRENT APPLICATION NUMBER: US/10/084,817
; CURRENT FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 60/270,784
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 365
; SOFTWARE: PERL Program
; SEQ ID NO 272
; LENGTH: 1271
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030119009A1 2818482CB1
US-10-084-817-272

Query Match      98.1%; Score 83.4; DB 5; Length 1271;
Best Local Similarity 98.8%; Pred. No. 4e-20;
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 60
DB 1160 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 1219

QY 61 CTGGTGCCCCCACTAAGTAGAATAA 85
DB 1220 CTGGTGCCCCCACTAAGTAGAATAA 1244

RESULT 5
US-10-159-563-162
Sequence 162, Application US/10159563
Publication No. US20040009154A1
GENERAL INFORMATION:
; APPLICANT: Khan, Javed
; APPLICANT: Ringner, Markus
; APPLICANT: Peterson, Carsten
; APPLICANT: Meltzer, Paul
; TITLE OF INVENTION: SELECTIONS OF GENES AND METHODS OF USING THE SAME FOR
; TITLE OF INVENTION: DIAGNOSIS AND FOR TARGETING THE THERAPY OF SELECT CANCERS
; FILE REFERENCE: 11613.56US11
; CURRENT APPLICATION NUMBER: US/10/159,563
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: US 10/133,937
; PRIOR FILING DATE: 2002-04-25
; NUMBER OF SEQ ID NOS: 444
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 162
; LENGTH: 1320
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-159-563-162

Query Match      98.1%; Score 83.4; DB 6; Length 1320;
Best Local Similarity 98.8%; Pred. No. 4e-20;
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 60
DB 1198 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 1257

QY 61 CTGGTGCCCCCACTAAGTAGAATAA 85
DB 1258 CTGGTGCCCCCACTAAGTAGAATAA 1282

RESULT 6
US-10-741-600-162
Sequence 162, Application US/10741600
Publication No. US20050026169A1
GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
```

```
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 162
; LENGTH: 1539
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-600-162

Query Match          98.1%; Score 83.4; DB 8; Length 1539;
Best Local Similarity 98.8%; Pred. No. 4.2e-20;
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 60
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1435 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 1494

Qy 61 CTGGGTGCCCCCACTAAGTAGAATAA 85
    ||||||||||||||||||||||||||||
Db 1495 CTGGGTGCCCCCACTAAGTAGAATAA 1519

RESULT 7
US-10-198-846-10394
; Sequence 10394, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10394
; LENGTH: 2027
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; LOCATION: 1804, 2025, 2026, 2027
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-10394

Query Match          98.1%; Score 83.4; DB 5; Length 2027;
Best Local Similarity 98.8%; Pred. No. 4.6e-20;
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 60
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1282 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 1341

Qy 61 CTGGGTGCCCCCACTAAGTAGAATAA 85
    ||||||||||||||||||||||||||||
Db 1342 CTGGGTGCCCCCACTAAGTAGAATAA 1366

RESULT 8
US-10-084-817-258
; Sequence 258, Application US/10084817
; Publication No. US20030119009A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
```

```
; APPLICANT: Susan Stuart
; APPLICANT: Jed G. Nuchtern
; APPLICANT: Sharon E. Plon
; APPLICANT: Jason M. Shohet
; TITLE OF INVENTION: GENES REGULATED BY MYCN ACTIVATION
; FILE REFERENCE: PA-0046 US
; CURRENT APPLICATION NUMBER: US/10/084,817
; CURRENT FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 60/270,784
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 365
; SOFTWARE: PERL Program
; SEQ ID NO 258
; LENGTH: 3674
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030119009A1 391741.16
US-10-084-817-258

Query Match          98.1%; Score 83.4; DB 5; Length 3674;
Best Local Similarity 98.8%; Pred. No. 5.5e-20;
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 60
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3501 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 3560

Qy 61 CTGGGTGCCCCCACTAAGTAGAATAA 85
    ||||||||||||||||||||||||||||
Db 3561 CTGGGTGCCCCCACTAAGTAGAATAA 3585

RESULT 9
US-10-741-600-17596
; Sequence 17596, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17596
; LENGTH: 16175
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-600-17596

Query Match          98.1%; Score 83.4; DB 8; Length 16175;
Best Local Similarity 98.8%; Pred. No. 8.9e-20;
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 60
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 10072 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 10131

Qy 61 CTGGGTGCCCCCACTAAGTAGAATAA 85
    ||||||||||||||||||||||||||||
Db 10132 CTGGGTGCCCCCACTAAGTAGAATAA 10156

RESULT 10
US-10-741-600-17595
; Sequence 17595, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
```

```
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001499
; CURRENT APPLICATION NUMBER: US/10/741.600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17595
; LENGTH: 20773
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-600-17595

Query Match          98.1%; Score 83.4; DB 8; Length 20773;
Best Local Similarity 98.8%; Pred. No. 9.6e-20;
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 60
Db 19163 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 19222

Qy 61 CTGGGTGCCCCACTAAGTAGAAAAA 85
Db 19223 CTGGGTGCCCCACTAAGTAGAATAA 19247

RESULT 11
US-10-741-600-17894
; Sequence 17894, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: C1001499
; CURRENT APPLICATION NUMBER: US/10/741.600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17894
; LENGTH: 28789
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-600-17894

Query Match          98.1%; Score 83.4; DB 8; Length 28789;
Best Local Similarity 98.8%; Pred. No. 1.1e-19;
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 60
Db 3965 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 4024

Qy 61 CTGGGTGCCCCACTAAGTAGAAAAA 85
Db 4025 CTGGGTGCCCCACTAAGTAGAATAA 4049

RESULT 12
US-09-880-107-3768
; Sequence 3768, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880.107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
```

```
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3768
; LENGTH: 66109
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 X66401
; NAME/KEY: unaure
; LOCATION: (1)..(66109)
; OTHER INFORMATION: n = a or c or g or t
US-09-880-107-3768

Query Match          98.1%; Score 83.4; DB 3; Length 66109;
Best Local Similarity 98.8%; Pred. No. 1.4e-19;
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 60
Db 37840 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 37899

Qy 61 CTGGGTGCCCCACTAAGTAGAAAAA 85
Db 37900 CTGGGTGCCCCACTAAGTAGAATAA 37924

RESULT 13
US-09-873-367C-1060
; Sequence 1060, Application US/09873367C
; Publication No. US20030165839A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; APPLICANT: Soppet, Daniel
; APPLICANT: Endress, Gregory
; APPLICANT: Augustus, Meena
; APPLICANT: Ebner, Reinhard
; APPLICANT: Carter, Kenneth
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
; FILE REFERENCE: 689290-64
; CURRENT APPLICATION NUMBER: US/09/873.367C
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: U.S. 60/236,891
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: U.S. 60/236,842
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: U.S. 60/244,867
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: U.S. 60/245,084
; PRIOR FILING DATE: 2000-11-01
; NUMBER OF SEQ ID NOS: 1067
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1060
; LENGTH: 66109
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(66109)
; OTHER INFORMATION: n=a,t,g or c
US-09-873-367C-1060

Query Match          98.1%; Score 83.4; DB 3; Length 66109;
Best Local Similarity 98.8%; Pred. No. 1.4e-19;
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 60
Db 37840 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 37899

Qy 61 CTGGGTGCCCCACTAAGTAGAAAAA 85
```

;  
CURRENT FILING DATE:

This Page Blank (usplg)



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 21:43:57 ; Search time 342.427 Seconds  
(without alignments)  
4099.008 Million cell updates/sec

Title: US-09-980-046B-9  
Perfect score: 30  
Sequence: 1 cgggaataaaggctgtgtaagacaaaaa 30  
Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 41078325 seqs, 23393541228 residues  
Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST: \*  
1: gb\_est1: \*  
2: gb\_est2: \*  
3: gb\_est3: \*  
4: gb\_hc: \*  
5: gb\_hc2: \*  
6: gb\_est4: \*  
7: gb\_est5: \*  
8: gb\_est6: \*  
9: gb\_est7: \*  
10: gb\_gaet1: \*  
11: gb\_gaet2: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30	100.0	123	6	CD520437 AGENCOURT
2	30	100.0	688	2	BG713896 602673757
3	28.4	94.7	110	3	BM055401 ie94d09.y
4	28.4	94.7	157	2	BI052127 PM0-GN040
5	28.4	94.7	166	1	AI401069 th22a01.x
6	28.4	94.7	203	1	AI342093 q898h01.x
7	28.4	94.7	241	5	BUS35679 AGENCOURT
8	28.4	94.7	245	6	CA426383 UI-H-FEI-
9	28.4	94.7	273	2	BF447271 7p47a11.x
10	28.4	94.7	274	1	AI400897 th30f11.x
11	28.4	94.7	299	3	BQ013504 UI-1-BC1p
12	28.4	94.7	300	3	BQ011929 UI-1-BC1p
13	28.4	94.7	327	1	AA133211 2k97g08.s
14	28.4	94.7	332	6	CF529736 UI-1-BC1p
15	28.4	94.7	349	5	BUS57786 AGENCOURT
16	28.4	94.7	377	5	BUI93594 AGENCOURT
17	28.4	94.7	380	3	BQ028373 UI-H-CO0-
18	28.4	94.7	412	1	AI952448 wx73h02.x
19	28.4	94.7	420	1	AW117898 xe38e06.x
20	28.4	94.7	432	1	AI336331 qt43f06.x
21	28.4	94.7	432	6	CB052146 NISC gl06
22	28.4	94.7	433	1	AI086993 oz64h09.x

23	28.4	94.7	451	1	AJ708390
24	28.4	94.7	457	5	BUI174042
c 25	28.4	94.7	465	2	BE856455
c 26	28.4	94.7	507	1	AW080368
c 27	28.4	94.7	542	3	BM999267
28	28.4	94.7	544	5	BUS96408
29	28.4	94.7	556	5	BUS21125
c 30	28.4	94.7	571	1	AA587336
c 31	28.4	94.7	575	1	AA149615
c 32	28.4	94.7	577	1	AA165422
c 33	28.4	94.7	584	1	AA583402
c 34	28.4	94.7	592	1	AAW001496
c 35	28.4	94.7	595	1	AA587899
c 36	28.4	94.7	598	1	AA439384
c 37	28.4	94.7	605	7	CR992231
c 38	28.4	94.7	618	5	BUI734305
c 39	28.4	94.7	625	5	BUI737708
c 40	28.4	94.7	629	5	BUE21270
41	28.4	94.7	629	7	CR999119
42	28.4	94.7	632	5	BQ946113
43	28.4	94.7	633	5	BUS96376
44	28.4	94.7	637	5	BQ437232
45	28.4	94.7	637	5	BUI84485

ALIGNMENTS

RESULT 1  
CD520437  
LOCUS  
DEFINITION AGENCOURT 14354662 NIH MGC 191 Homo sapiens cDNA clone  
IMAGE: 30411401 5', mRNA sequence.  
ACCESSION CD520437  
VERSION CD520437.1  
KEYWORDS GI:31452155  
SOURCE EST.  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE 1 (bases 1 to 123)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Daniela S. Gerhard, Ph.D.  
Office of Cancer Genomics  
National Cancer Institute / NIH  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
Tissue Procurement: Narayan Bhat  
CDNA Library Preparation: CLONTECH Laboratories, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Cloned by: Agencourt Bioscience Corporation  
Cloned distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: NUCM203 row: b column: 18  
High quality sequence stop: 123.  
Location/Qualifiers  
1. 123  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:30411401"  
/tissue\_type="Pooled"  
/lab\_host="DH10B (TI phage-resistant)"  
/clone\_lib="NIH\_MGC\_191"  
/note="Vector: pDNR-LIB; Site 1: Sfil (ggccattatggcc) ; Site 2: Sfil (ggcgcctctggcc); Library is oligo-dT primed and directionally cloned. PBMC - Peripheral Blood Mononuclear Cells. RNA was pooled from 3/6hour stimulation with PMA adn Ionomycin. 5' and 3' adaptors were used in

cloning as follows: 5' adaptor sequence:  
5'-CAGGCCATTATGCC-3' and 3' adaptor sequence:  
5'-ATTCTAGAGCGCGGCGGACATG-dT(30)BN-3' (where B = A,  
C, or G and N = A, C, G, or T). Average insert size 1.69  
kb (range 0.70-5.0 kb). 15/15 colonies contained inserts  
by PCR. This library was enriched for full-length clones  
and was constructed by Clontech Laboratories (Palo Alto,  
CA). Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 100.0%; Score 30; DB 6; Length 123;  
Best Local Similarity 100.0%; Pred. No. 0.63;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGAATAAAGGCTGTTGTAAGACAAAA 30  
Db 63 CGGAATAAAGGCTGTTGTAAGACAAAA 92

RESULT 2  
LOCUS BG713896 688 bp mRNA linear EST 08-MAY-2001  
DEFINITION 602673757F1 NIH\_MGC\_96 Homo sapiens cDNA clone IMAGE:4796078 5',  
mRNA sequence.

ACCESSION BG713896  
VERSION BG713896.1 GI:13992827  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.

1 (bases 1 to 688)

NIH-MGC http://mgi.nci.nih.gov/.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.

Email: cga@bse@mail.nih.gov  
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki  
Toshiyuki and Piero Carninci (RIKEN)

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov  
Plate: L1AM10680 row: b column: 15  
High quality sequence stop: 664.

## FEATURES

Location/Qualifiers  
1..688  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4796078"  
/tissue\_type="hypothalamus"  
/lab\_host="DH10B"  
/clone\_lib="NIH\_MGC\_96"

/notes="Organ: Brain; Vector: pBluescriptR (modified  
pBluescript KS+); Site\_1: BamHI; Site\_2: SalI-XhoI  
(gtcgag); Oligo-dT primed using primer  
5'-TTTTTTTTTTTTTTVN-3', size-selected for average  
insert size 2.3 kb and normalized to ROT 5. This is a  
primary library enriched for full-length clones and  
constructed using the Cap-trapper method (Carninci, in  
preparation). Library constructed by M. Brownstein  
(NIH/NHGRI, National Institutes of Health). Note: this is  
a NIH\_MGC Library."

## ORIGIN

Query Match 100.0%; Score 30; DB 2; Length 688;  
Best Local Similarity 100.0%; Pred. No. 0.69;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGAATAAAGGCTGTTGTAAGACAAAA 30  
Db 637 CGGAATAAAGGCTGTTGTAAGACAAAA 666

## RESULT 3

LOCUS BM055401

DEFINITION

BM055401

VERSION

BM055401.1

KEYWORDS

EST.

SOURCE

ORGANISM

Homo sapiens

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 110)

Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,

Lemishka, I., Searce, M., Brestelli, J., Gradwohl, G., Clifton, S.,

Hillier, L., Marra, M., Pape, D., Wyllie, T., Martin, J., Blistain, A.,

Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J.,

Cardenas, M., Gibbons, M., McCann, R., Cole, R., Tsagarishvili, R.,

Williams, T., Jackson, Y. and Bowers, Y.

Endocrine Pancreas Consortium

Unpublished (2000)

Other ESTs: i94d09.x1

Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue

Endocrine Pancreas Consortium

Harvard University, Howard Hughes Medical Institute

Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,

MA 02138

Tel: 617-495-1812

Fax: 617-495-8557

Email: dmelton@biohp.harvard.edu

Library was constructed by Dr. Douglas Melton DNA sequencing by:

Washington University Genome Sequencing Center For information on

obtaining a clone please contact: Juliana Brown

(brown@fas.harvard.edu) This sequence now available from the IMAGE

consortium, for clone orders contact: info@image.llnl.gov.

Location/Qualifiers

1..110

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:5674433"

/sex="Both"

/tissue\_type="Islets of Langerhans"

/dev\_stage="Adult"

/lab\_host="DH10B"

/clone\_lib="Melton Normalized Human Islet 4 N4-HIS 1"

/note="Organ: Pancreas; Vector: pSPORT1; Site\_1: Not 1;

Site\_2: Sal 1; Starting library constructed using

SuperScript Plasmid Library kit (Life Technologies). cDNA

made by oligo-dT priming. Size-selected by column

fractionation; average insert size 1.08 kb. Library was

amplified once on solid support and plasmid DNA from

library was prepared. The library DNA was normalized by

method #4 from Bonaldo, Lennon, and Soares 1996 Genome

Research 6:791-806; 0.5 microgram single-stranded library

plasmid DNA was mixed with 5 micrograms PCR product

representing library inserts and hybridized to an Ecot of

20. Single-stranded (unhybridized) plasmids were isolated

by hydroxyapatite chromatography and used to make this

library."

## ORIGIN

Query Match 94.7%; Score 28.4; DB 3; Length 110;  
Best Local Similarity 96.7%; Pred. No. 2.5;  
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGAATAAAGGCTGTTGTAAGACAAAA 30

```

Db      80  CGGAATAAAGCGCTGTTGTAAGAGAAAA 109
|||||
RESULT 4
BIOS2127
LOCUS   157 bp  mRNA  linear  EST 15-JUN-2001
DEFINITION PMO-GN0405-180101-001-e04 GN0405 Homo sapiens cDNA, mRNA sequence.
ACCESSION BI052127
VERSION   BI052127.1 GI:14459657
KEYWORDS EST.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
          Hominoidea; Homo.
REFERENCE 1 (bases 1 to 157)
AUTHORS  Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
          Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.P.,
          Goldman,G.H., Carvalho,A.F., Matukuma,A., Baia,G.S., Simpson,D.H.,
          Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
          O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.P., de Souza,S.J. and
          Simpson,A.J.
          Shotgun sequencing of the human transcriptome with ORF expressed
          sequence tags
          Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
          10737800
          Contact: Simpson A.J.G.
          Laboratory of Cancer Genetics
          Ludwig Institute for Cancer Research
          Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
          Brazil
          Tel: +55-11-2704922
          Fax: +55-11-2707001
          Email: asimpson@ludwig.org.br
          This sequence was derived from the FAPESP/LICR Human Cancer Genome
          Project. This entry can be seen in the following URL
          (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=PM0&t2=PM0-GN0405-
          180101-001-e04&t3=2001-01-18&t4=1)
          Seq primer: puc 18 forward
          High quality sequence start: 2
          High quality sequence stop: 156.
          Location/Qualifiers
            1..157
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /dev_stage="Adult"
              /clone_lib="GN0405"
              /note="Organ: placenta normal; Vector: puc18; Site 1:
              Sma1; Site 2: Sma1; A mini-library was made by cloning
              products derived from ORESTES PCR (U.S. Letters Patent
              application No. 196,716 - Ludwig Institute for Cancer
              Research) profiles into the puc 18 vector. Reverse
              transcription of tissue mRNA and cDNA amplification were
              performed under low stringency conditions."
          ORIGIN
            Query Match 94.7%; Score 28.4; DB 2; Length 157;
            Best Local Similarity 96.7%; Pred. No. 2.6;
            Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1  CGGAATAAAGCGCTGTTGTAAGAGAAAA 30
|||||
Db      123 CGGAATAAAGCGCTGTTGTAAGAGAAAA 152
|||||
RESULT 5
AI401069/c
LOCUS   166 bp  mRNA  linear  EST 08-FEB-1999
DEFINITION th22a01.x1 NCI CGAP Pr28 Homo sapiens cDNA clone IMAGE:2118984 3',
          similar to gb:MI3755 INTERFERON-INDUCED 17 KD PROTEIN (HUMAN);,
          mRNA sequence.
          REFERENCE
          AUTHORS  NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
          National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
          Tumor Gene Index
          Unpublished (1997)

```

```

ACCESSION AI401069
VERSION   AI401069.1 GI:4244156
KEYWORDS EST.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
          Hominoidea; Homo.
REFERENCE 1 (bases 1 to 166)
AUTHORS  NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
          National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
          Tumor Gene Index
          Unpublished (1997)
JOURNAL
COMMENT   Contact: Robert Strausberg, Ph.D.
          Email: cgapsb@mail.nih.gov
          Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.
          Emmert-Buck, M.D., Ph.D.
          CDNA Library Preparation: M. Bento Soares, Ph.D.
          CDNA Library Arrayed by: Greg Lennon, Ph.D.
          DNA sequencing by: Washington University Genome Sequencing Center
          Clone distribution: NCI-CGAP clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          www-bio.llnl.gov/bbrp/image/image.html
          Seq primer: -40UP from Gibco.
          Location/Qualifiers
            1..166
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="IMAGE:2118984"
              /sex="male"
              /dev_stage="adult"
              /lab_host="DH10B"
              /clone_lib="NCI CGAP Pr28"
              /note="Organ: prostate; Vector: p773D-Pac (Pharmacia)
              with a modified polylinker; Plasmid DNA from the
              normalized library NCI CGAP Pr22 was prepared, and ss
              circles were made in vitro. Following HAP purification,
              this DNA was used as tracer in a subtractive hybridization
              reaction. The driver was PCR-amplified cDNAs from a pool
              of 5,000 clones made from the same library (cloneIDs
              985608-986759, 1101192-1101959, and 1217928-1220615).
              Subtraction by Bento Soares and M. Fatima Bonaudo. "
          ORIGIN
            Query Match 94.7%; Score 28.4; DB 1; Length 166;
            Best Local Similarity 96.7%; Pred. No. 2.6;
            Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1  CGGAATAAAGCGCTGTTGTAAGAGAAAA 30
|||||
Db      44  CGGAATAAAGCGCTGTTGTAAGAGAAAA 15
|||||
RESULT 6
AI342093/c
LOCUS   203 bp  mRNA  linear  EST 29-DEC-1998
DEFINITION q998h01.x1 NCI CGAP GC4 Homo sapiens cDNA clone IMAGE:1946161 3',
          similar to gb:MI3755 INTERFERON-INDUCED 17 KD PROTEIN (HUMAN);,
          mRNA sequence.
          ACCESSION AI342093
          VERSION   AI342093.1 GI:4079020
          KEYWORDS EST.
          SOURCE   Homo sapiens (human)
          ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
          Hominoidea; Homo.
          REFERENCE 1 (bases 1 to 203)
          AUTHORS  NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
          National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
          Tumor Gene Index
          Unpublished (1997)

```

COMMENT Contact: Robert Strauberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: M. Bento Soares, Ph.D.  
cDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone Distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html  
Seq primer: -40UP from Gibco.

## FEATURES

source

1. .203  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:1946161"  
/issue\_type="pooled germ cell tumors"  
/lab\_host="DH10B"  
/clone\_lib="NCI CGAP GC4"  
/notes="Vector: pT7T3D-Fac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from 3 pooled germ cell tumors, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library is normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo."

## ORIGIN

Query Match 94.7%; Score 28.4; DB 1; Length 203;  
Best Local Similarity 96.7%; Pred. No. 2.6;  
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGAATAAAGGCTGTTGTTAAAGACAAAAA 30  
|||||  
Db 52 CGGAATAAAGGCTGTTGTTAAAGACAAAAA 23  
|||||

RESULT 7  
BU535679  
LOCUS AGENCOURT 10217871 NIH MGC 107 Homo sapiens cDNA clone  
DEFINITION IMAGE:6563371 5', mRNA sequence.  
ACCESSION BU535679  
VERSION BU535679.1 GI:22846120  
KEYWORDS EST.  
SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.

## REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 241)  
NIH-MGC http://mgs.nci.nih.gov/.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strauberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov  
Plate: LLCM2740 row: o column: 19  
High quality sequence stop: 240.

## FEATURES

source

1. .241  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6563371"

/tissue\_type="adenocarcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_107"  
/notes="Organ: breast; Vector: pOTB7; Site\_1: EcoRI;  
Site\_2: XhoI; cDNA made by oligo-dT priming.  
Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match

Best Local Similarity

Matches

Conservative

Mismatches

Indels

Gaps

0;

Qy

Db

RESULT 8

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 245)

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strauberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: James Martin

cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Clone distribution information can be obtained

from Dr. M. Bento Soares, bento-soares@uiowa.edu

Seq primer: M13 FORWARD

POLYA=Yes.

Location/Qualifiers

1. .245

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="UI-H-FEI-beh-g-13-0-UI"

/tissue\_type="Cell lines"

/dev\_stage="Adult"

/lab\_host="DH10B (Life Technologies)"

/clone\_lib="NCI CGAP FEI"

/notes="Organ: chondrosarcoma; Vector: pT7T3-Pac

(Pharmacia) with a modified polylinker. Site\_1: EcoRI;

Site\_2: Not I; NCI CGAP FEI is a normalized cDNA library

derived from a pool of mRNA obtained from 3 cell lines

from grade II chondrosarcoma tissues. The library was

constructed according to Bonaldo, Lennon and Soares,

Genome Research, 6:791-806, 1996. First strand cDNA

synthesis was primed with an oligo-dT primer containing a

Not I site. Double stranded cDNA was ligated to an EcoRI

adaptor, digested with Not I, and cloned directionally

into pT7T3-Pac vector. The oligonucleotide used to prime

the synthesis of first-strand cDNA contains a library tag

sequence that is located between the Not I site and the

|||||  
|||  
2 CGGAATTAAGGCTGTTCATAAGCAAAA 3  
  
AI400897      274 bp    mRNA      linear      EST 30-MAR-1999  
th30fil.x1 NCI CQAP Pr28 Homo sapiens cDNA clone IMAGE:2119821 3'  
similar to gb:M13755 INTERFERON-INDUCED 17 KD PROTEIN (HUMAN);,  
mRNA sequence.

similar to gb:M13755.1  
mRNA sequence.

ACCESSION A1400897  
VERSION A1400897.1  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

KEYWORDS	EST.	ORGANISM
SOURCE	Homo sapiens (human)	
	Homo sapiens	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
	Mammalia; Euthera; Euarchontoglires; Primates; Catarrhini;	
	Hominidae; Homo.	
REFERENCE	1 (bases 1 to 274)	

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

1 (bases 1 to 274)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D.  
cDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html  
Insert Length: 359 Std Error: 0.00  
Seg primer: -40UP from Gibco.

Seq primer: -40UP from Gibco.
Location/Qualifiers
1. .274
FEATURES
source

```
/organism="Homo sapiens"  
/mol_type="mRNA"  
/db_xref="taxon:9606"
```

```
/db_xref="taxon:9606"
/clone="IMAGE:2119821"
/sex="male"
```

```
/sex="male"
/dev_stage="adult"
/lab_host="DH10B"
```

```
/lab_host="DH10B"  
/clone_lib="NCI_CGAP_Pr28"  
/note="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia)
```

/note="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Plasmid DNA from the normalized library NCI CGAP Pr22 was prepared, and ss

normalized library NCI CGAP Pr22 was prepared, and ss circles were made *in vitro*. Following HAP purification, this DNA was used as tracer in a subtractive hybridization

reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs

of 5,000 clones made from the same library (clones 985608-986759, 1101192-1101959, and 1217928-1220615). Subtraction by Bento Soares and M. Fatima Bonaldo. "

ORIGIN

Subtraction by Bento Soares and M. Fatima Bonalquo.

Query Match	94.7%;	Score 28.4;	DB 1;	Length 274;
Best Local Similarity	96.7%;	Pred. No. 2.7;		

```

0-:
1-:
2-:
3-:
4-:
5-:
6-:
7-:
8-:
9-:
10-:
11-:
12-:
13-:
14-:
15-:
16-:
17-:
18-:
19-:
20-:
21-:
22-:
23-:
24-:
25-:
26-:
27-:
28-:
29-:
30-:
31-:
32-:
33-:
34-:
35-:
36-:
37-:
38-:
39-:
40-:
41-:
42-:
43-:
44-:
45-:
46-:
47-:
48-:
49-:
50-:
51-:
52-:
53-:
54-:
55-:
56-:
57-:
58-:
59-:
60-:
61-:
62-:
63-:
64-:
65-:
66-:
67-:
68-:
69-:
70-:
71-:
72-:
73-:
74-:
75-:
76-:
77-:
78-:
79-:
80-:
81-:
82-:
83-:
84-:
85-:
86-:
87-:
88-:
89-:
90-:
91-:
92-:
93-:
94-:
95-:
96-:
97-:
98-:
99-:
100-:
101-:
102-:
103-:
104-:
105-:
106-:
107-:
108-:
109-:
110-:
111-:
112-:
113-:
114-:
115-:
116-:
117-:
118-:
119-:
120-:
121-:
122-:
123-:
124-:
125-:
126-:
127-:
128-:
129-:
130-:
131-:
132-:
133-:
134-:
135-:
136-:
137-:
138-:
139-:
140-:
141-:
142-:
143-:
144-:
145-:
146-:
147-:
148-:
149-:
150-:
151-:
152-:
153-:
154-:
155-:
156-:
157-:
158-:
159-:
160-:
161-:
162-:
163-:
164-:
165-:
166-:
167-:
168-:
169-:
170-:
171-:
172-:
173-:
174-:
175-:
176-:
177-:
178-:
179-:
180-:
181-:
182-:
183-:
184-:
185-:
186-:
187-:
188-:
189-:
190-:
191-:
192-:
193-:
194-:
195-:
196-:
197-:
198-:
199-:
200-:
201-:
202-:
203-:
204-:
205-:
206-:
207-:
208-:
209-:
210-:
211-:
212-:
213-:
214-:
215-:
216-:
217-:
218-:
219-:
220-:
221-:
222-:
223-:
224-:
225-:
226-:
227-:
228-:
229-:
230-:
231-:
232-:
233-:
234-:
235-:
236-:
237-:
238-:
239-:
240-:
241-:
242-:
243-:
244-:
245-:
246-:
247-:
248-:
249-:
250-:
251-:
252-:
253-:
254-:
255-:
256-:
257-:
258-:
259-:
260-:
261-:
262-:
263-:
264-:
265-:
266-:
267-:
268-:
269-:
270-:
271-:
272-:
273-:
274-:
275-:
276-:
277-:
278-:
279-:
280-:
281-:
282-:
283-:
284-:
285-:
286-:
287-:
288-:
289-:
290-:
291-:
292-:
293-:
294-:
295-:
296-:
297-:
298-:
299-:
300-:
301-:
302-:
303-:
304-:
305-:
306-:
307-:
308-:
309-:
310-:
311-:
312-:
313-:
314-:
315-:
316-:
317-:
318-:
319-:
320-:
321-:
322-:
323-:
324-:
325-:
326-:
327-:
328-:
329-:
330-:
331-:
332-:
333-:
334-:
335-:
336-:
337-:
338-:
339-:
340-:
341-:
342-:
343-:
344-:
345-:
346-:
347-:
348-:
349-:
350-:
351-:
352-:
353-:
354-:
355-:
356-:
357-:
358-:
359-:
360-:
361-:
362-:
363-:
364-:
365-:
366-:
367-:
368-:
369-:
370-:
371-:
372-:
373-:
374-:
375-:
376-:
377-:
378-:
379-:
380-:
381-:
382-:
383-:
384-:
385-:
386-:
387-:
388-:
389-:
390-:
391-:
392-:
393-:
394-:
395-:
396-:
397-:
398-:
399-:
400-:
401-:
402-:
403-:
404-:
405-:
406-:
407-:
408-:
409-:
410-:
411-:
412-:
413-:
414-:
415-:
416-:
417-:
418-:
419-:
420-:
421-:
422-:
423-:
424-:
425-:
426-:
427-:
428-:
429-:
430-:
431-:
432-:
433-:
434-:
435-:
436-:
437-:
438-:
439-:
440-:
441-:
442-:
443-:
444-:
445-:
446-:
447-:
448-:
449-:
450-:
451-:
452-:
453-:
454-:
455-:
456-:
457-:
458-:
459-:
460-:
461-:
462-:
463-:
464-:
465-:
466-:
467-:
468-:
469-:
470-:
471-:
472-:
473-:
474-:
475-:
476-:
477-:
478-:
479-:
480-:
481-:
482-:
483-:
484-:
485-:
486-:
487-:
488-:
489-:
490-:
491-:
492-:
493-:
494-:
495-:
496-:
497-:
498-:
499-:
500-:
501-:
502-:
503-:
504-:
505-:
506-:
507-:
508-:
509-:
510-:
511-:
512-:
513-:
514-:
515-:
516-:
517-:
518-:
519-:
520-:
521-:
522-:
523-:
524-:
525-:
526-:
527-:
528-:
529-:
530-:
531-:
532-:
533-:
534-:
535-:
536-:
537-:
538-:
539-:
540-:
541-:
542-:
543-:
544-:
545-:
546-:
547-:
548-:
549-:
550-:
551-:
552-:
553-:
554-:
555-:
556-:
557-:
558-:
559-:
560-:
561-:
562-:
563-:
564-:
565-:
566-:
567-:
568-:
569-:
570-:
571-:
572-:
573-:
574-:
575-:
576-:
577-:
578-:
579-:
580-:
581-:
582-:
583-:
584-:
585-:
586-:
587-:
588-:
589-:
590-:
591-:
592-:
593-:
594-:
595-:
596-:
597-:
598-:
599-:
600-:
601-:
602-:
603-:
604-:
605-:
606-:
607-:
608-:
609-:
610-:
611-:
612-:
613-:
614-:
615-:
616-:
617-:
618-:
619-:
620-:
621-:
622-:
623-:
624-:
625-:
626-:
627-:
628-:
629-:
630-:
631-:
632-:
633-:
634-:
635-:
636-:
637-:
638-:
639-:
640-:
641-:
642-:
643-:
644-:
645-:
646-:
647-:
648-:
649-:
650-:
651-:
652-:
653-:
654-:
655-:
656-:
657-:
658-:
659-:
660-:
661-:
662-:
663-:
664-:
665-:
666-:
667-:
668-:
669-:
670-:
671-:
672-:
673-:
674-:
675-:
676-:
677-:
678-:
679-:
680-:
681-:
682-:
683-:
684-:
685-:
686-:
687-:
688-:
689-:
690-:
691-:
692-:
693-:
694-:
695-:
696-:
697-:
698-:
699-:
700
```

Qy 1 CGGAATATAAGCTGTTTAAAGACAAAA 30  
|||  
Dh 22 CCGCATTATTAAGCTCTTTGTAAAGACAAA 4

Db 33 CGGAATAAAGGCTGTTGTTAAAGAGAAAA 4

RESULT 11  
R0013504/C

BQ013504/c
LOCUS
BQ013504          299 bp       linear     EST 26-MAR-2002
DEFINITION
UI-1-BC1b-akr-a-03-0-UI.s1 NCI CGAP pl3 Homo sapiens cDNA clone

DEFINITION UI-1-BC1p-akr-g-03-0-UI.s1 NCI CGAP pl3 Homo sapiens cDNA clone  
UI-1-BC1p-akr-g-03-0-UI 3', mRNA sequence.  
ACCESSION B0013504

ACCESSION BQ013504  
VERSION BQ013504.1 GI:19738405  
KEYWORDS EST.

**KEYWORDS**      **EST.**

**SOURCE**  
**ORGANISM** Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.  
**REFERENCE**  
**AUTHORS** 1 (bases 1 to 299)  
**TITLE** NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
**COMMENT** National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
 Unpublished (1997)  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue Procurement: Dr. Steven Brown  
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Clone distribution information can be obtained from Dr. M. Bento Soares, [bento-soares@uiowa.edu](mailto:bento-soares@uiowa.edu)  
 Seq primer: M13 FORWARD  
 POLYA=Yes.

**FEATURES**  
**source**  
 Location/Qualifiers  
 1..299  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="UI-1-BC1p-akr-g-03-0-UI"  
 /tissue\_type="Placenta"  
 /dev\_stage="8-9 weeks"  
 /lab\_host="DH10B (Life Technologies)"  
 /clone\_lib="NCI CGAP P13"  
 /note="Organ: Placenta; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP P13 is a subtracted cDNA library constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tags for this library are GA, AGGAA. For additional information, contact: Bento Soares, [bento-soares@uiowa.edu](mailto:bento-soares@uiowa.edu)  
 TAG\_TISSUE=placenta human 8 week  
 TAG\_LIB=UI-1-BC1p  
 TAG\_SEQ=GA"

**ORIGIN**  
 Query Match 94.7%; Score 28.4; DB 3; Length 299;  
 Best Local Similarity 96.7%; Pred. No. 2.7;  
 Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 CGGAATAAAGGCTGTTGTTAAAGACAAAA 30  
 ||||||||||||||||||||||||||||||||  
 Db 40 CGGAATAAAGGCTGTTGTTAAAGACAAAA 11  
 ||||||||||||||||||||||||||||||||

**RESULT 12**  
**LOCUS** BQ011929/c  
**DEFINITION** UI-1-BC1p-ati-b-04-0-UI.s1 NCI CGAP P13 Homo sapiens cDNA clone  
**ACCESSION** BQ011929  
**VERSION** BQ011929.1 GI:19736830  
**KEYWORDS** EST.  
**SOURCE** Homo sapiens (human)  
**ORGANISM** Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.  
**REFERENCE**  
**AUTHORS** 1 (bases 1 to 300)  
 NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

**TITLE**  
**JOURNAL**  
**COMMENT** National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
 Unpublished (1997)  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue Procurement: Dr. Steven Brown  
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Clone distribution information can be obtained from Dr. M. Bento Soares, [bento-soares@uiowa.edu](mailto:bento-soares@uiowa.edu)  
 Seq primer: M13 FORWARD  
 POLYA=Yes.

**FEATURES**  
**source**  
 Location/Qualifiers  
 1..300  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="UI-1-BC1p-ati-b-04-0-UI"  
 /tissue\_type="Placenta"  
 /dev\_stage="8-9 weeks"  
 /lab\_host="DH10B (Life Technologies)"  
 /clone\_lib="NCI CGAP P13"  
 /note="Organ: Placenta; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP P13 is a subtracted cDNA library constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tags for this library are GA, AGGAA. For additional information, contact: Bento Soares, [bento-soares@uiowa.edu](mailto:bento-soares@uiowa.edu)  
 TAG\_TISSUE=placenta human 8 week  
 TAG\_LIB=UI-1-BC1p  
 TAG\_SEQ=GA"

**ORIGIN**  
 Query Match 94.7%; Score 28.4; DB 3; Length 300;  
 Best Local Similarity 96.7%; Pred. No. 2.7;  
 Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 CGGAATAAAGGCTGTTGTTAAAGACAAAA 30  
 ||||||||||||||||||||||||||||||||  
 Db 41 CGGAATAAAGGCTGTTGTTAAAGACAAAA 12  
 ||||||||||||||||||||||||||||||||

**RESULT 13**  
**LOCUS** AAI33211/c  
**DEFINITION** AAI33211/c 327 bp mRNA linear EST 14-MAY-1997  
 IMAGE:490814.3' similar to gb:U13755 INTERFERON-INDUCED 17 KD PROTEIN (HUMAN); mRNA sequence.  
**ACCESSION** AAI33211  
**VERSION** AAI33211.1 GI:1689973  
**KEYWORDS** EST.  
**SOURCE** Homo sapiens (human)  
**ORGANISM** Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.  
**REFERENCE**  
**AUTHORS** 1 (bases 1 to 327)  
 Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiatelli, B., Chisoe, S., Dietrich, N., DuBuque, T., Favello, A., Gish, W., Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N., Rohlfing, T., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Schellberg, K., Soares, M.B., Tan, F., Thiermy-Meg, J., Trevaskis, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R. and Marra, M.



```
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_107"
/notes="Organ: breast; Vector: pOTB7; Site_1: EcoRI;
Site_2: XhoI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
```

ORIGIN

```
Query Match      94.7%; Score 28.4; DB 5; Length 349;
Best Local Similarity 96.7%; Pred. No. 2.7;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1  CGGAATAAAGGCTGTTGTTAAAGACAAAA 30
      |||||||
Db      285 CGGAATAAAGGCTGTTGTTAAAGACAAAA 314
```

Search completed: December 7, 2005, 09:56:03  
Job time : 345.527 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 10:04:41 ; Search time 44.9561 Seconds  
(without alignments)  
4072.612 Million cell updates/sec

Title: US-09-980-046B-8

Perfect score: 103

Sequence: 1 cggacagacggcagcagctcc.....tgagcccttctcacaaaa 103

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA:\*

- 1: /cgn2\_6/pdata/1/ina/1 COMB.seq:\*
- 2: /cgn2\_6/pdata/1/ina/5 COMB.seq:\*
- 3: /cgn2\_6/pdata/1/ina/6A COMB.seq:\*
- 4: /cgn2\_6/pdata/1/ina/6B COMB.seq:\*
- 5: /cgn2\_6/pdata/1/ina/H COMB.seq:\*
- 6: /cgn2\_6/pdata/1/ina/PCTUS COMB.seq:\*
- 7: /cgn2\_6/pdata/1/ina/PP COMB.seq:\*
- 8: /cgn2\_6/pdata/1/ina/RE COMB.seq:\*
- 9: /cgn2\_6/pdata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	101.4	98.4	427	3	US-09-461-325-107
2	101.4	98.4	427	3	US-10-012-542-107
3	101.4	98.4	427	3	US-10-115-123-107
4	100.4	97.5	1277	3	US-09-461-325-20
5	100.4	97.5	1277	3	US-10-012-542-20
6	100.4	97.5	1277	3	US-10-115-123-20
7	74.4	72.2	280	3	US-09-118-554-58
8	74.4	72.2	280	3	US-09-118-627-58
9	74.4	72.2	280	3	US-09-602-877A-58
10	31.4	30.5	4118	2	US-08-119-125A-3
11	28.6	27.8	601	3	US-09-949-016-87323
12	28.6	27.8	32393	3	US-09-949-016-14203
13	28	27.2	160759	3	US-09-949-016-16514
14	27.6	26.8	1387	3	US-09-663-600A-56
15	27.6	26.8	29636	3	US-09-949-016-12189
16	27.6	26.8	29637	3	US-09-949-016-17173
17	27.6	26.8	29638	3	US-09-949-016-14137
18	27.6	26.8	135667	3	US-09-949-016-15051
19	27.6	26.8	152486	3	US-09-949-016-12869
20	27.4	26.6	601	3	US-09-949-016-147021
21	27.4	26.6	374159	3	US-09-949-016-15091
22	27	26.2	193169	3	US-09-949-016-15091
23	26.8	26.0	113538	3	US-09-949-016-16329
24	26.8	26.0	246230	3	US-09-949-016-17019

25	26.8	26.0	246230	3	US-09-949-016-17020	Sequence 17020, A
26	26.8	26.0	246230	3	US-09-949-016-17021	Sequence 17021, A
27	26.8	26.0	246230	3	US-09-949-016-17022	Sequence 17022, A
c 28	26.6	25.8	2260	3	US-09-976-594-91	Sequence 91, Appl
c 29	26.4	25.6	601	3	US-09-949-016-52888	Sequence 52888, A
30	26.4	25.6	31440	3	US-09-949-016-12578	Sequence 12578, A
31	26.4	25.6	31444	3	US-09-949-016-16400	Sequence 16400, A
32	26.4	25.6	90876	3	US-09-949-016-13271	Sequence 13271, A
33	26.4	25.6	152331	3	US-09-128-155-16	Sequence 16, Appl
34	26.4	25.6	176373	3	US-09-128-155-17	Sequence 17, Appl
c 35	26.2	25.4	601	3	US-09-949-016-40554	Sequence 40554, A
c 36	26.2	25.4	601	3	US-09-949-016-194471	Sequence 194471, A
37	26.2	25.4	450395	3	US-09-949-016-17588	Sequence 17588, A
38	26	25.2	34078	3	US-09-949-016-17588	Sequence 765, App
39	26	25.2	129554	3	US-09-949-002-765	Sequence 3859, App
40	25.8	25.0	406	3	US-09-533-559-3859	Sequence 3859, App
c 41	25.8	25.0	448	3	US-09-513-999C-35605	Sequence 35605, A
42	25.8	25.0	2949	3	US-09-942-711-24	Sequence 24, Appl
43	25.8	25.0	2958	3	US-09-942-711-23	Sequence 23, Appl
44	25.8	25.0	2958	3	US-09-942-711-25	Sequence 25, Appl
45	25.8	25.0	2978	3	US-09-942-711-21	Sequence 21, Appl

ALIGNMENTS

RESULT 1  
US-09-461-325-107  
Sequence 107, Application US/09461325A  
Patent No. 6475753  
GENERAL INFORMATION:  
APPLICANT: Ruben et al.  
TITLE OF INVENTION: 9A Human Secreted Proteins  
FILE REFERENCE: P2028P1  
CURRENT APPLICATION NUMBER: US/09/461,325A  
CURRENT FILING DATE: 1999-12-14  
EARLIER APPLICATION NUMBER: PCT/US99/13418  
EARLIER FILING DATE: 1999-06-15  
EARLIER APPLICATION NUMBER: 60/089,507  
EARLIER FILING DATE: 1998-06-16  
EARLIER APPLICATION NUMBER: 60/089,508  
EARLIER FILING DATE: 1998-06-16  
EARLIER APPLICATION NUMBER: 60/089,509  
EARLIER FILING DATE: 1998-06-16  
EARLIER APPLICATION NUMBER: 60/089,510  
EARLIER FILING DATE: 1998-06-16  
EARLIER APPLICATION NUMBER: 60/090,112  
EARLIER FILING DATE: 1998-06-22  
EARLIER APPLICATION NUMBER: 60/090,113  
EARLIER FILING DATE: 1998-06-22  
NUMBER OF SEQ ID NOS: 532  
SOFTWARE: Patentin ver. 2.0  
SEQ ID NO 107  
LENGTH: 427  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-461-325-107

Query Match 98.4%; Score 101.4; DB 3; Length 427;

Best Local Similarity 99.0%; Pred. No. 1.6e-28;

Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGACAGCGGCGAGTCCAGCTCTGGTTCCCTCGGTTTATCTGTAGATGAA 60  
Db 281 CGGACAGCGGCGAGTCCAGCTCTGGTTCCCTCGGTTTATCTGTAGATGAA 340

Qy 61 ATGGTTCCTCCATAATAAGGGCGCATGAGCCCTTCCTCAAAAA 103

Db 341 ATGGTTCCTCCATAATAAGGGCGCATGAGCCCTTCCTCAAAAA 383

RESULT 2

US-10-012-542-107

```
; Sequence 107, Application US/10012542
; Patent No. 662741
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 94 Human Secreted Proteins
; FILE REFERENCE: P2029P1
; CURRENT APPLICATION NUMBER: US/10/012,542
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/461,325
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-12-14
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,507
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,508
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,509
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,510
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,112
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,113
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 107
; LENGTH: 427
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-012-542-107

Query Match          98.4%; Score 101.4; DB 3; Length 427;
Best Local Similarity 99.0%; Pred. No. 1.6e-28;
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGACAGACGGCAGCAGTCCAGCTCTGGTTTCCTTCGCGTTTATTCCTGTTAGAAATGAA 60
Db 281 CGGACAGACGGCAGCAGTCCAGCTCTGGTTTCCTTCGCGTTTATTCCTGTTAGAAATGAA 340

Qy 61 ATGGTTCCCATAAATAAGGGGCATGAGCCCTTCCTCACAATAA 103
Db 341 ATGGTTCCCATAAATAAGGGGCATGAGCCCTTCCTCACAATAA 383

RESULT 4
US-09-461-325-20
; Sequence 20, Application US/09461325A
; Patent No. 6475753
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 94 Human Secreted Proteins
; FILE REFERENCE: P2029P1
; CURRENT APPLICATION NUMBER: US/09/461,325A
; CURRENT FILING DATE: 1999-12-14
; EARLIER APPLICATION NUMBER: PCT/US99/13418
; EARLIER FILING DATE: 1999-06-15
; EARLIER APPLICATION NUMBER: 60/089,507
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/089,508
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/089,509
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/089,510
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/090,112
; EARLIER FILING DATE: 1998-06-22
; EARLIER APPLICATION NUMBER: 60/090,113
; EARLIER FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 1277
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1207)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1272)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-461-325-20

Query Match          97.5%; Score 100.4; DB 3; Length 1277;
Best Local Similarity 98.1%; Pred. No. 6.2e-28;
Matches 101; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGACAGACGGCAGCAGTCCAGCTCTGGTTTCCTTCGCGTTTATTCCTGTTAGAAATGAA 60
Db 1130 CGGACAGACGGCAGCAGTCCAGCTCTGGTTTCCTTCGCGTTTATTCCTGTTAGAAATGAA 1189

Qy 61 ATGGTTCCCATAAATAAGGGGCATGAGCCCTTCCTCACAATAA 103
Db 1190 ATGGTTCCCATAAATAAGGGGCATGAGCCCTTCCTCACAATAA 1232

RESULT 5
US-10-012-542-20
; Sequence 20, Application US/10012542
; Patent No. 662741
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 94 Human Secreted Proteins
; FILE REFERENCE: P2029P1
; CURRENT APPLICATION NUMBER: US/10/012,542
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/461,325
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-12-14
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,507
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,508
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,509
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,510
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,112
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,113
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 107
; LENGTH: 427
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-012-542-107

Query Match          98.4%; Score 101.4; DB 3; Length 427;
Best Local Similarity 99.0%; Pred. No. 1.6e-28;
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGACAGACGGCAGCAGTCCAGCTCTGGTTTCCTTCGCGTTTATTCCTGTTAGAAATGAA 60
Db 281 CGGACAGACGGCAGCAGTCCAGCTCTGGTTTCCTTCGCGTTTATTCCTGTTAGAAATGAA 340

Qy 61 ATGGTTCCCATAAATAAGGGGCATGAGCCCTTCCTCACAATAA 103
Db 341 ATGGTTCCCATAAATAAGGGGCATGAGCCCTTCCTCACAATAA 383

RESULT 3
US-10-115-123-107
; Sequence 107, Application US/10115123
; Patent No. 6774216
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 94 Human Secreted Proteins
; FILE REFERENCE: P2029G30AP1D2
; CURRENT APPLICATION NUMBER: US/10/115,123
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: PCT/US99/13418
; PRIOR FILING DATE: 1999-06-15
; PRIOR APPLICATION NUMBER: 60/089,507
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089,508
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089,509
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089,510
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/090,112
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090,113
; PRIOR FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 107
; LENGTH: 427
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-115-123-107
```



```
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 58
; LENGTH: 280
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-118-627-58

Query Match      72.2%; Score 74.4; DB 3; Length 280;
Best Local Similarity 98.7%; Pred. No. 2.5e-18;
Matches 75; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 25 TCTGGTTTCCTTCGCGTTTATTCTGTAGAAATGGTTCCTCCATAAATAGGGGCAT 84
Db 266 TCTGGTTTCCTTCGCGTTTATTCTGTAGAAATGGTTCCTCCATAAATAGGGGCAT 207

Qy 85 GAGCCCTTCCTCAAA 100
Db 206 GAGCCCTTCCTCAGCA 191

RESULT 9
US-09-602-877A-58/c
; Sequence 58, Application US/09602877A
; Patent No. 6432707
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.446C5
; CURRENT APPLICATION NUMBER: US/09/602,877A
; CURRENT FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 58
; LENGTH: 280
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-602-877A-58

Query Match      72.2%; Score 74.4; DB 3; Length 280;
Best Local Similarity 98.7%; Pred. No. 2.5e-18;
Matches 75; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 25 TCTGGTTTCCTTCGCGTTTATTCTGTAGAAATGGTTCCTCCATAAATAGGGGCAT 84
Db 266 TCTGGTTTCCTTCGCGTTTATTCTGTAGAAATGGTTCCTCCATAAATAGGGGCAT 207

Qy 85 GAGCCCTTCCTCAAA 100
Db 206 GAGCCCTTCCTCAGCA 191

RESULT 10
US-08-119-125A-3/c
; Sequence 3, Application US/08119125A
; Patent No. 5610011
; GENERAL INFORMATION:
; APPLICANT: SMITH, Hilda Elizabeth
; APPLICANT: VECHT, Uri
; TITLE OF INVENTION: DNA Sequences which code for Virulence
; TITLE OF INVENTION: Characteristics of Streptococcus suis and parts thereof, poly
; TITLE OF INVENTION: antibodies derived therefrom and the use thereof for the diagn
; TITLE OF INVENTION: protection against infection by S. suis in mammals, including
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Centraal Diergeneeskundig Instituut
; STREET: Bdelherweg 15
; CITY: PH Lelystad
; STATE:
; COUNTRY: The Netherlands
```

```
; ZIP: NL-8219
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
; OPERATING SYSTEM: MS-DOS v.6.0
; SOFTWARE: WordPerfect v. 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/119,125A
; FILING DATE: 20-SEP-1993
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/NL92/00054
; FILING DATE: 19-MAR-1992
; APPLICATION NUMBER: NL 9100510
; FILING DATE: 21-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Handal, Anthony H.
; REGISTRATION NUMBER: 26275
; REFERENCE/DOCKET NUMBER: SMITHHE119125
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (203) 838-8589
; TELEFAX: (203) 838-8794
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4118 base pairs
; TYPE: Nucleic acid with corresponding amino acids
; STRANDEDNESS: single stranded
; TOPOLOGY: linear
; MOLECULE TYPE: genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus suis type II (pathogenic)
; FEATURE:
; OTHER INFORMATION: Muramidase released protein (MRP) gene
; FEATURE:
; NAME/KEY: promoter -35 region
; LOCATION: bp 4 to 9
; FEATURE:
; NAME/KEY: promoter -10 region
; LOCATION: bp 29 to 34
; FEATURE:
; NAME/KEY: promoter -35 region
; LOCATION: bp 40 to 45
; FEATURE:
; NAME/KEY: promoter -10 region
; LOCATION: bp 63 to 68
; FEATURE:
; NAME/KEY: ribosome binding site
; LOCATION: bp 147 to 152
; FEATURE:
; NAME/KEY: signal peptide
; LOCATION: bp 159 to 299
; FEATURE:
; NAME/KEY: mature peptide
; LOCATION: bp 300 to 3926
; FEATURE:
; NAME/KEY: proline rich region
; LOCATION: from bp 2757 to 3014
; FEATURE:
; NAME/KEY: repetitive units
; LOCATION: from bp 3015 to 3176, 3423 to 3584 and 3585 to 3743
; FEATURE:
; NAME/KEY: membrane anchor sequence
; LOCATION: from bp 3825 to 3926
; FEATURE:
; NAME/KEY: dyad symmetry regions
; LOCATION: from bp 4069 to 4080 and from bp 4087 to 4098
; US-08-119-125A-3

Query Match      30.5%; Score 31.4; DB 2; Length 4118;
Best Local Similarity 64.4%; Pred. No. 0.19;
Matches 47; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

Qy 29 GTTTCCTTCGCGTTTATTCTGTAGAAATGGTTCCTCCATAAATAGGGGCATGAGC 88
```

Db 3778 GTTTCCTCTCAGGTGTTGTTGGTTATTAGAAAATTTCGATAAATATGTCGTGTTA 3719

Qy 89 CCTTCCTCACAA 101

Db 3718 CCTTCCTCATCAA 3706

RESULT 11  
US-09-949-016-87323  
? Sequence 87323, Application US/09949016  
? Patent No. 6812339  
? GENERAL INFORMATION:  
? APPLICANT: VENTER, J. Craig et al.  
? TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
? TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
? FILE REFERENCE: CL001307  
? CURRENT APPLICATION NUMBER: US/09/949,016  
? CURRENT FILING DATE: 2000-04-14  
? PRIOR APPLICATION NUMBER: 60/241,755  
? PRIOR FILING DATE: 2000-10-20  
? PRIOR APPLICATION NUMBER: 60/237,768  
? PRIOR FILING DATE: 2000-10-03  
? PRIOR APPLICATION NUMBER: 60/231,498  
? PRIOR FILING DATE: 2000-09-08  
? NUMBER OF SEQ ID NOS: 207012  
? SOFTWARE: FastSeq for Windows Version 4.0  
? SEQ ID NO 87323  
? LENGTH: 601  
? TYPE: DNA  
? ORGANISM: Human  
US-09-949-016-87323

RESULT 12  
US-09-949-016-14203/c  
; Sequence 14203, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 14203  
; LENGTH: 32393  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-14203

Db 5562 GTCACTGATCTGCTTCTCTTGGCTTTAGGAGGCTCTGATTCAAATAACT 5504

RESULT 13  
US-09-949-016-16514  
; Sequence 16514, Application US/09949016  
; Patent No. 5812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 16514  
; LENGTH: 160759  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(160759)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-16514

RESULT 14  
US-09-663-600A-56  
; Sequence 56, Application US/09663600A  
; Patent No. 6573068  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste  
; APPLICANT: Duclert, Aymeric  
; APPLICANT: Bougueleret, Lydie  
; TITLE OF INVENTION: EXTENDED CDNAS FOR SECRETED PROTEINS  
; FILE REFERENCE: 31.US3.CIP  
; CURRENT APPLICATION NUMBER: US/09/663,600A  
; CURRENT FILING DATE: 2000-09-15  
; PRIOR APPLICATION NUMBER: 09/191,997  
; PRIOR FILING DATE: 1998-11-13  
; PRIOR APPLICATION NUMBER: 60/066,677  
; PRIOR FILING DATE: 1997-11-13  
; PRIOR APPLICATION NUMBER: 60/069,957  
; PRIOR FILING DATE: 1997-12-17  
; PRIOR APPLICATION NUMBER: 60/074,121  
; PRIOR FILING DATE: 1998-02-09  
; PRIOR APPLICATION NUMBER: 60/081,563  
; PRIOR FILING DATE: 1998-04-13  
; PRIOR APPLICATION NUMBER: 60/096,116  
; PRIOR FILING DATE: 1998-08-10  
; PRIOR APPLICATION NUMBER: 60/099,273  
; PRIOR FILING DATE: 1998-09-04  
; NUMBER OF SEQ ID NOS: 229  
; SOFTWARE: Patent.pm





***This Page Blank (uspto)***



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 11:47:47 ; Search time 462.191 Seconds  
(without alignments)  
4239.411 Million cell updates/sec

Title: US-09-980-046B-7

Perfect score: 294

Sequence: 1 cggcgatcgccgcgcttat.....agcgagcacctagacaaaa 294

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N\_Geneseq\_21.\*

- 1: geneseqn1980s.\*
- 2: geneseqn1990s.\*
- 3: geneseqn2000s.\*
- 4: geneseqn2001as.\*
- 5: geneseqn2001bs.\*
- 6: geneseqn2002as.\*
- 7: geneseqn2002bs.\*
- 8: geneseqn2003as.\*
- 9: geneseqn2003bs.\*
- 10: geneseqn2003cs.\*
- 11: geneseqn2003ds.\*
- 12: geneseqn2004as.\*
- 13: geneseqn2004bs.\*
- 14: geneseqn2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	294	100.0	294	4	AAC89738 Human gas
2	269.4	91.6	2369	2	Aav71777 Human mem
3	269.4	91.6	2369	3	Aa292428 cDNA enco
4	269.4	91.6	2741	13	ADR07477
5	269.4	91.6	3047	5	AAS81031
6	267.8	91.1	603	6	ABV94671 Human pan
7	267.8	91.1	3079	12	ADN03645 Antipsori
8	267.8	91.1	3079	13	ACN38103
9	267.8	91.1	3079	13	ADT04107
10	253.6	86.3	452	9	ACH47801 Human inf
11	244.4	83.1	485	9	ACH29871 Human tes
12	102	34.7	466	13	ADQ51104 Novel can
13	35.8	12.2	35973	6	ABK13076 Human amy
14	32.4	11.0	1559	13	ADR31515 Human ade
15	32.4	11.0	2724	6	ABZ12539 Arabidops
16	32.2	11.0	2000	8	ADA71938 Rice gene
17	32.2	11.0	120239	13	ABD33244 Murine ca
18	32	10.9	6499	4	AAK86748 Human imm
19	31.6	10.7	105219	11	ACN44286 Human gen

c 20	31.4	10.7	602	12	ADQ19063
c 21	31.4	10.7	2000	12	ADJ40696
c 22	31.4	10.7	3931	5	AAS29195
c 23	31.4	10.7	3931	6	ABS68335
c 24	31.4	10.7	3931	10	ADC25457
c 25	31.4	10.7	50784	11	ACN45090_3
c 26	31.2	10.6	127722	12	ADQ97301
c 27	31	10.5	5259	13	ADR06908
c 28	31	10.5	81656	12	ADQ97876
c 29	30.8	10.5	567	5	AAS64324
c 30	30.8	10.5	1809	11	ACH99470
c 31	30.8	10.5	2111	5	ABV27899
c 32	30.8	10.5	2111	5	ABV22063
c 33	30.8	10.5	3223	4	ABL08588
c 34	30.8	10.5	42999	8	ABX11086
c 35	30.6	10.4	238	10	ACD92285
c 36	30.6	10.4	238	10	ACD92063
c 37	30.6	10.4	238	10	ACD92284
c 38	30.6	10.4	238	10	ACD92154
c 39	30.6	10.4	238	10	ACD92286
c 40	30.6	10.4	238	10	ACD92228
c 41	30.6	10.4	238	10	ACD92062
c 42	30.6	10.4	239	10	ACD92041
c 43	30.6	10.4	260	12	ADQ05366
c 44	30.6	10.4	277	10	ACD92096
c 45	30.6	10.4	277	10	ACD92097

#### ALIGNMENTS

RESULT 1

AAC89738

ID AAC89738 standard; cDNA; 294 BP.

XX AAC89738;

XX 12-MAR-2001 (first entry)

XX Human gastrointestinal inflammation-related cDNA, SEQ ID NO: 7.

XX Human; cytostatic; immunomodulator; immunostimulant; vulnery;

XX anti-inflammatory; neuroprotective; antibacterial; gene therapy;

XX gastrointestinal inflammation; immune system disorder; genetic disorder;

XX cancer; autoimmune disorder; infection; wound healing; ss.

XX Homo sapiens.

XX WO200073324-A2.

XX 07-DEC-2000.

XX 01-JUN-2000; 2000WO-US015191.

XX 01-JUN-1999; 99US-0137058P.

XX (DIGI-) DIGITAL GENE TECHNOLOGIES INC.

XX Youakim A, Dubose RF, Sims JE, Pribyl TM, Hillbush BS, Haseel KW;

XX WPI; 2001-061508/07.

XX New polynucleotides and polypeptides, useful in gene therapy and in

XX diagnosing a pathological condition, e.g. for modulating gene expression

XX in gastrointestinal inflammation, or for treating cancers or genetic

XX disorders.

XX Claim 1; Page 85; 108pp; English.

XX The present sequence is one of a number of isolated human polynucleotides

XX which are useful in gene therapy, and for diagnosing a pathological

XX condition or a susceptibility to it. In particular, the polynucleotides

XX are useful for modulating gene expression in gastrointestinal

CC inflammation. The polynucleotides are useful for chromosome  
 CC identification, controlling gene expression through triple helix  
 CC formation or antisense DNA or RNA, or identifying individuals from minute  
 CC biological samples using DNA-based identification techniques. The  
 CC polynucleotides can also be used as an alternative to restriction  
 CC fragment length polymorphism (RFLP), by determining the actual base-by-  
 CC base DNA sequences of selected portions of an individual's genome. The  
 CC polynucleotides may also be used as molecular weight markers on Southern  
 CC gels, as diagnostic probes for the presence of a specific mRNA, as a  
 CC probe to substract-out known sequences in the process of discovering novel  
 CC polynucleotides, or as an antigen to elicit an immune response. The  
 CC polypeptides are useful in diagnostic procedures to detect a disorder.  
 CC The polynucleotides and polypeptides are useful for preventing, treating  
 CC or ameliorating immune system disorders, genetic disorders, cancers, some  
 CC autoimmune disorders, or infections. The polynucleotides and polypeptides  
 CC are also useful for differentiating, proliferating or attracting cells,  
 CC leading to the regeneration of tissues, especially in wounds or burns.  
 CC The polypeptides and polynucleotides may also be used as a food additive  
 CC or preservative

XX  
 SQ Sequence 294 BP; 75 A; 87 C; 71 G; 61 T; 0 U; 0 Other;

Query Match 100.0%; Score 294; DB 4; Length 294;  
 Best Local Similarity 100.0%; Pred. No. 4.6e-88;  
 Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CGGCGATGCGGCGGCTTATCCCATGATGCTCTGTCAAGGAGTGACATACCTACT 60  
 DB 1 CGGCGATGCGGCGGCTTATCCCATGATGCTCTGTCAAGGAGTGACATACCTACT 60  
 QY 61 GAGGAGCCCCACCCGCCCTCTGATGACCGAATCAGGAGTAGAGTTAACTGCAGCG 120  
 DB 61 GAGGAGCCCCACCCGCCCTCTGATGACCGAATCAGGAGTAGAGTTAACTGCAGCG 120  
 QY 121 GAACATGTCTATTTCTCTATTTCTGTGACGCGCCGCCATGCCCCACCCACCAAGAGG 180  
 DB 121 GAACATGTCTATTTCTCTATTTCTGTGACGCGCCGCCATGCCCCACCCACCAAGAGG 180  
 QY 181 CAGGAGGCGCCAGTCATGACCTCTCTGTGAGGGGTCTGTACAGGTTCTTCTGAAAG 240  
 DB 181 CAGGAGGCGCCAGTCATGACCTCTCTGTGAGGGGTCTGTACAGGTTCTTCTGAAAG 240  
 QY 241 CTTTATAGCAGTAAACAGTTTGACATGAAATAAAGCGAGCAGCCTAGACAAAAA 294  
 DB 241 CTTTATAGCAGTAAACAGTTTGACATGAAATAAAGCGAGCAGCCTAGACAAAAA 294

RESULT 2  
 ID AAV7177 standard; DNA; 2369 BP.  
 XX  
 AC AAV7177;  
 XX

09-FEB-1999 (first entry)

XX Human membrane fusion protein WDPPro2 encoding DNA.

XX WDPPro1; WDPPro2; WD-40; membrane fusion protein; recombinant; cancer;  
 KW beta-transducin; signal transduction; developmental disorder; ss.

XX Homo sapiens.

XX Key Location/Qualifiers  
 FT CDS 119..1939  
 FT /\*tag= a  
 FT /product= "WDPPro2"

US5846777-A.

08-DEC-1998.

26-JUN-1997; 97US-00883534.

PR 26-JUN-1997; 97US-00883534.  
 XX (INCY-) INCYTE PHARM INC.

XX Lal P, Bandman O, Corley NC;

XX WPI; 1999-059062/05.

DR P-PSDB; AAW84086.

XX DNA encoding WD-40 proteins - useful for producing recombinant proteins.

XX Claim 10; Fig 3A-F; 42pp; English.

XX This DNA encodes a human membrane fusion protein WDPPro2. The invention  
 CC provides polynucleotide sequences (AAV7177 and AAV71777) encoding human  
 CC membrane fusion proteins WDPPro1 and WDPPro2. Host cells transformed with  
 CC vectors containing the sequences can be used producing recombinant  
 CC proteins. WD-40 proteins (collectively designated WDPPro) are homologues  
 CC of beta-transducin and function in signal transduction pathways within  
 CC the cytoplasm. The WDPPro proteins may be used in the diagnosis,  
 CC prevention, and treatment of cancers and developmental disorders

XX SQ Sequence 2369 BP; 560 A; 660 C; 695 G; 454 T; 0 U; 0 Other;

Query Match 91.6%; Score 269.4; DB 2; Length 2369;

Best Local Similarity 99.6%; Pred. No. 2e-79;

Matches 270; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 21 TCCCATGATGCTCTCTGTCAAGGAGTGAGCAATCACCTACTGAGGAGGCCCGCCCT 80

DB 1898 TCCCATGATGCTCTCTGTCAAGGAGTGAGCAATCACCTACTGAGGAGGCCCGCCCT 1957

QY 81 CTGGATGGACCGAATCAGGAGCTAGAGTTAACTGACGCGAATGTCATTTCTCTATT 140

DB 1958 CTGGATGGACCGAATCAGGAGCTAGAGTTAACTGACGCGAATGTCATTTCTCTATT 2017

QY 141 TCTGTGACGCGGCCCATGCCCCACCCACCAAGAGGAGGAGGCCCGCCCTATGAC 200

DB 2018 TCTGTGACGCGGCCCATGCCCCACCCACCAAGAGGAGGAGGCCCGCCCTATGAC 2077

QY 201 CCTCGCTCTGTGACGGGTGTCTGTACACGTTCTTCTGAAAGCTTTAGACAGTAACAGTTTG 260

DB 2078 CCTCGCTCTGTGACGGGTGTCTGTACACGTTCTTCTGAAAGCTTTAGACAGTAACAGTTTG 2137

QY 261 CACATGAAAAATAAGCGGAGCCTAGACNA 291

DB 2138 CACATGAAAAATAAGCGGAGCCTAGACNA 2168

RESULT 3

AAZ92428

ID AAZ92428 standard; cDNA; 2369 BP.

XX AAZ92428;

XX 05-JUN-2000 (first entry)

XX cDNA encoding human WD-40 protein, WDPPro2.

XX WDPPro2; human; neutrophil; WD-40 protein; beta-transducin homologue;  
 KW signal transduction pathway; cytoplasmic; cancer; developmental disorder;  
 KW prevention; treatment; diagnosis; drug screening; ss.

XX Homo sapiens.

XX Key Location/Qualifiers  
 FT CDS 119..1939  
 FT /\*tag= a  
 FT /product= "Human WDPPro2"

US6025464-A.

15-FEB-2000.

```
XX 02-DEC-1998; 98US-00204764.
XX
XX 26-JUN-1997; 97US-00883534.
XX
XX (INCY-) INCYTE PHARM INC.
XX
XX Lal P, Corley NC, Bandman O;
XX
XX WPI; 2000-181836/16.
XX
XX P-PSDB; AAY77489.
XX
XX New WD-40 protein used for the diagnosis, prevention and treatment of
XX diseases and conditions associated with WD-40 expression e.g. cancer and
XX developmental disorders.
XX
XX Example 5; Fig 3A-F; 42pp; English.
XX
XX The present sequence represents cDNA encoding a human WD-40 protein,
XX WDPro2. cDNAs encoding WD-Pro2 were isolated from a neurophil cDNA
XX library, this sequence being a consensus. Chemical and structural
XX homology exists between WDPro1 and WD-40 protein (GI 1384131). WD-40
XX proteins are homologues of beta-transducin and function in signal
XX transduction pathways within the cytoplasm. In addition, WDPro2 is
XX expressed in cancers and transformed cells, indicating that it may play a
XX role in cancer and developmental disorders, particularly disorders in
XX which WDPro2 is overexpressed. WDPro2 and WDPro1 (AAV77488) may be used
XX for the prevention, treatment and diagnosis of cancer and developmental
XX disorders e.g., cerebral palsy, or Crohn's disease. The proteins can be
XX used to generate antibodies which can be used to screen libraries of
XX pharmaceutical agents to identify antagonists or agonists of WD-40
XX protein activity
XX
XX Sequence 2369 BP; 560 A; 560 C; 595 G; 454 T; 0 U; 0 Other;
XX
Query Match 91.6%; Score 269.4; DB 3; Length 2369;
Best Local Similarity 99.6%; Pred. No. 2e-79;
Matches 270; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 21 TCCCATGATGCTCTGTCAAGAGTGGACATCACTACTGAGGAGGCCACCCCGCCT 80
Db 1898 TCCCATGATGCTCTGTCAAGAGTGGACATCACTACTGAGGAGGCCACCCCGCCT 1957
Qy 81 CTGATGACCGAATCAGGAGTGGAGTTTAACTGCGAGCGGAACATGTCATTCTCTATT 140
Db 1958 CTGATGACCGAATCAGGAGTGGAGTTTAACTGCGAGCGGAACATGTCATTCTCTATT 2017
Qy 141 TCTGTGACGCGCCCCCATGCCCCACCCACCAAGAGGCGAGGCGCCAGTCATGAC 200
Db 2018 TCTGTGACGCGCCCCCATGCCCCACCCACCAAGAGGCGAGGCGCCAGTCATGAC 2077
Qy 201 CCTCGTCTCTGACGGGTGTCGTACAGTTCCTTCTGAAAGCTTTAGACAGTAACAGTTTG 260
Db 2078 CCTCGTCTCTGACGGGTGTCGTACAGTTCCTTCTGAAAGCTTTAGACAGTAACAGTTTG 2137
Qy 261 CACATGAAAAATAAAGCGAGCAGCTAGACAA 291
Db 2138 CACATGAAAAATAAAGCGAGCAGCTAGACAA 2168
RESULT 4
ADRO7477
ID ADRO7477 standard; cDNA; 2741 BP.
XX
XX AC ADRO7477;
XX
XX 04-NOV-2004 (first entry)
XX
XX Full length human cDNA useful for treating neurological disease Seq 983.
XX
XX gene; ss; human; oligo-capping method; diagnostic marker; gene therapy;
XX osteoporosis; neurological disease; Alzheimer's disease;
XX Parkinson's disease; dementia; short memory; cancer;
KW sense or motor function; emotional reaction; fear response; panic;
KW osteopathic; neuroprotective; nootropic; antiparkinsonian; cyostatic;
KW tranquiliser.
XX
XX Homo sapiens.
XX
XX EP1447413-A2.
XX
XX 18-AUG-2004.
XX
XX 12-FEB-2004; 2004EP-00003145.
XX
XX 14-FEB-2003; 2003JP-00102207.
XX
XX 09-MAY-2003; 2003JP-00131452.
XX
XX (REAS-) RES ASSOC BIO TECHNOLOGY.
XX
XX Isogai T, Yamamoto J, Nishikawa T, Isono Y, Sugiyama T, Otsuki T;
XX Wakamatsu A, Ishii S, Nagai K, Irie R;
XX
XX WPI; 2004-583265/57.
XX
XX P-PSDB; ADR09433.
XX
XX New 1995 cDNA, useful for treating osteoporosis, neurological diseases,
XX Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.
XX
XX Claim 1; SEQ ID NO 983; 2686pp; English.
XX
XX This invention relates to novel, isolated full length human cDNA
XX molecules and the encoded proteins thereof. Specifically, it refers to
XX cDNA clones obtained by an oligo-capping method, where none of these
XX clones are identical to any known human mRNAs. The present invention
XX describes an immunoassay to identify agonists and antagonists, as well as
XX antibodies, antisense molecules and siRNAs that can all be used to bind
XX to and modulate expression of the cDNA molecules. As such, these
XX molecules are useful for diagnostic markers or therapeutic targets for
XX the various diseases or morbid states. In particular, they are useful in
XX gene therapy for treating osteoporosis, neurological disease, Alzheimer's
XX disease, Parkinson's disease, dementia, short memory and various cancers,
XX as well as for maintaining equilibrium of sense or motor function, and
XX for treating emotional reaction, fear response and panic. Accordingly,
XX they exhibit osteopathic, neuroprotective, nootropic, antiparkinsonian,
XX cyostatic and tranquiliser activities. This polynucleotide is a full
XX length human cDNA sequence of the invention. NOTE: This sequence is not
XX given in the sequence listing of the specification but can be obtained on
XX CD-ROM from the European Patent Office, Vienna Sub-office.
XX
XX Sequence 2741 BP; 643 A; 744 C; 761 G; 593 T; 0 U; 0 Other;
XX
Query Match 91.6%; Score 269.4; DB 13; Length 2741;
Best Local Similarity 99.6%; Pred. No. 2.2e-79;
Matches 270; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 21 TCCCATGATGCTCTGTCAAGAGTGGACATCACTACTGAGGAGGCCACCCCGCCT 80
Db 1803 TCCCATGATGCTCTGTCAAGAGTGGACATCACTACTGAGGAGGCCACCCCGCCT 1862
Qy 81 CTGATGACCGAATCAGGAGTGGAGTTTAACTGCGAGCGGAACATGTCATTCTCTATT 140
Db 1863 CTGATGACCGAATCAGGAGTGGAGTTTAACTGCGAGCGGAACATGTCATTCTCTATT 1922
Qy 141 TCTGTGACGCGCCCCCATGCCCCACCCACCAAGAGGCGAGGCGCCAGTCATGAC 200
Db 1923 TCTGTGACGCGCCCCCATGCCCCACCCACCAAGAGGCGAGGCGCCAGTCATGAC 1982
Qy 201 CCTCGTCTCTGACGGGTGTCGTACAGTTCCTTCTGAAAGCTTTAGACAGTAACAGTTTG 260
Db 1983 CCTCGTCTCTGACGGGTGTCGTACAGTTCCTTCTGAAAGCTTTAGACAGTAACAGTTTG 2042
Qy 261 CACATGAAAAATAAAGCGAGCAGCTAGACAA 291
Db 2043 CACATGAAAAATAAAGCGAGCAGCTAGACAA 2073
```

```

RESULT 5
AAS81031
ID AAS81031 standard; cDNA; 3047 BP.
XX
AC AAS81031;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #16835.
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
XX Homo sapiens.
XX
XX WO200175067-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US008631.
XX
XX 31-MAR-2000; 2000US-00540217.
XX
XX 23-AUG-2000; 2000US-00649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
XX
XX P-PSDB; ABG16844.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity.
XX
XX Claim 1; SEQ ID NO 16835; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridisation probes, polymerase chain
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX and in recombinant production of (II). The polynucleotides are also used
XX in diagnostics as expressed sequence tags for identifying expressed
XX genes. (I) is useful in gene therapy techniques to restore normal
XX activity of (II) or to treat disease states involving (II). (II) is
XX useful for generating antibodies against it, detecting or quantitating a
XX polypeptide in tissue, as molecular weight markers and as a food
XX supplement. (II) and its binding partners are useful in medical imaging
XX of sites expressing (II). (I) and (II) are useful for treating disorders
XX involving aberrant protein expression or biological activity. The
XX polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. AAS84197-AAS94564 represent novel human diagnostic
XX coding sequences of the invention. Note: The sequence data for this
XX patent did not appear in the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 3047 BP; 714 A; 825 C; 851 G; 657 T; 0 U; 0 Other;
XX
XX Query Match 91.6%; Score 269.4; DB 5; Length 3047;
XX Best Local Similarity 99.6%; Pred. No. 2.3e-79;
XX Matches 270; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX 21 TCCCATGATGCTTGTCTCAAGGAGTGGACATCACTACTGAGGAGGCCACCCCGCCT 80
XX
XX 1989 TCCCATGATGCTTGTCTCAAGGAGTGGACATCACTACTGAGGAGGCCACCCCGCCT 2048
XX
XX 81 CTGGATGACCGAATCAGGAGCTAGAGTTTAACCTGCAGCGGAACATGCTCTCTATT 140
XX

```

---

```

Db 2049 CTGGATGACCGAATCAGGAGCTAGAGTTTAACCTGCAGCGGAACATGCTCTCTATT 2108
Qy 141 TCTGTGACGGCGCCCATGCCCCCAGCCACCCACCAAGAGCGAGGGCCAGTCATGAC 200
XX
Db 2109 TCTGTGACGGCGCCCATGCCCCCAGCCACCCACCAAGAGCGAGGGCCAGTCATGAC 2168
Qy 201 CCTCGTCTCTGCAGGGTGTCTGTACACAGTTTCTTGAAGCTTTTACAGAGTTTG 260
Db 2169 CCTCGTCTCTGCAGGGTGTCTGTACACAGTTTCTTGAAGCTTTTACAGAGTTTG 2228
Qy 261 CACATGAAAAATAAAGCGAGCACCTAGACAA 291
XX
Db 2229 CACATGAAAAATAAAGCGAGCACCTAGACAA 2259
XX
RESULT 6
ABV94671
ID ABV94671 standard; cDNA; 603 BP.
XX
AC ABV94671;
XX
XX 14-JAN-2003 (first entry)
XX
XX Human pancreatic cancer expressed cDNA SEQ ID NO 44.
XX
XX Human; pancreas; cancer; gene therapy; vaccine; immunostimulant;
XX KW cytostatic; tumour; gene; ss.
XX
XX Homo sapiens.
XX
XX WO200260317-A2.
XX
XX 08-AUG-2002.
XX
XX 30-JAN-2002; 2002WO-US002781.
XX
XX 30-JAN-2001; 2001US-0265305P.
XX
XX 31-JAN-2001; 2001US-0265882P.
XX
XX 09-FEB-2001; 2001US-0267568P.
XX
XX 21-MAR-2001; 2001US-0278651P.
XX
XX 28-APR-2001; 2001US-0287112P.
XX
XX 16-MAY-2001; 2001US-0291631P.
XX
XX 12-JUL-2001; 2001US-0305484P.
XX
XX 20-AUG-2001; 2001US-0313999P.
XX
XX 27-NOV-2001; 2001US-0333626P.
XX
XX (CORI-) CORIXA CORP.
XX
XX Benson DR, Kalos MD, Lodes MJ, Persing DH, Hepler WT, Jiang Y;
XX
XX WPI; 2002-627435/67.
XX
XX New isolated polynucleotide and pancreatic tumor polypeptides, useful for
XX diagnosing, preventing and/or treating cancer, particularly pancreatic
XX cancer.
XX
XX Claim 1; SEQ ID NO 44; 300pp + Sequence Listing; English.
XX
XX The invention relates to an isolated polynucleotide (I) comprising: (a)
XX any of a group of over 4000 nucleotide sequences (ABV94628-ABV99145); (b)
XX complements of (a); (c) sequences consisting of at least 20 contiguous
XX residues of (a); (d) sequences that hybridize to (a), under moderately
XX stringent conditions; (e) sequences having at least 75% or 90% identity
XX to (a); or (f) degenerate variants of (a). Polypeptides (ABP68596-
XX ABP68637) encoded by (I) and oligonucleotide can be used to detect cancer
XX in a patient and compositions comprising polypeptides, polynucleotides,
XX antibodies, fusion proteins, T cell populations and antigen presenting
XX cells expressing the polypeptide are useful in treating pancreatic cancer
XX and stimulating an immune response. The polynucleotides can be used as
XX probes or primers for nucleic acid hybridisation, in the design and
XX preparation of ribozyme molecules for inhibiting expression of the tumour
XX polypeptides and proteins in the tumour cells, in vaccines and for gene
XX therapy. Note: The sequence data for this patent did not form part of the

```





PA (DRMA/) DRMANAC R T.  
PA (LABA/) LABAT I.  
PA (STAC/) STACHE-CRAIN B.  
PA (DICK/) DICKSON M C.  
PA (JONE/) JONES L W.  
XX  
XX  
PI Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;  
XX  
XX WPI; 2003-615964/58.  
XX  
XX New polynucleotide sequences obtained from various cDNA libraries, useful  
PT as hybridization probes, as oligomers for PCR, for chromosome and gene  
PT mapping, in the recombinant production of protein, or in generating  
PT antisense DNA or RNA.  
XX  
XX Claim 1; SEQ ID NO 35013; 44pp; English.  
XX  
XX The invention relates to an isolated polynucleotide comprising any one of  
CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was  
CC determined by the technique of SBH (sequencing by hybridisation). Also  
CC included is a purified polypeptide comprising a sequence corresponding to  
CC a reading frame of the novel polynucleotide. The nucleic acid sequences  
CC are useful in diagnostics as expressed sequence tags (EST) for  
CC identifying expressed genes or for physical mapping of the human genome,  
CC in forensics, in assessing biodiversity, or in identifying mutations  
CC responsible for genetic disorders and other traits. The nucleotide  
CC sequences are also useful as hybridisation probes, as oligomers for PCR,  
CC for chromosome and gene mapping, in the recombinant production of  
CC protein, or in generating antisense DNA or RNA. The purified polypeptide  
CC is useful for generating antibodies specific for it. The present sequence  
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data  
CC for this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from USPTO at  
CC seqdata.uspto.gov/sequence.html?DocID=20030073623  
XX  
XX SQ Sequence 452 BP; 122 A; 131 C; 108 G; 91 T; 0 U; 0 Other;  
Query Match 86.3%; Score 253.6; DB 9; Length 452;  
Best Local Similarity 98.2%; Pred. No. 2e-74;  
Matches 267; Conservative 0; Mismatches 4; Indels 1; Gaps 1;  
QY 21 TCCCATGATGCTCTGTCAAGAGTGGACAATCACCTACTGAGGAGCCCAACCCCGCCT 80  
Db 129 TCCCATGATGCTCTGTCAAGAGTGGACAATCACCTACTGAGGAGCCCAACCCCGCCT 188  
QY 81 CTGGATGACCGCAATCAGGACTAGAGTTTAACTGACGGGACATGTCATTCTCTATT 140  
Db 189 CTGGATGACCGCAATCAGGACTAGAGTTTAACTGACGGGACATGTCATTCTCTATT 248  
QY 141 TCTGTGACGCGCCGCCCATGCCCCCACCACCAAGAGGAGGAGGCCCATGTCATGAC 200  
Db 249 TCTGTGACGCGCCGCCCATGCCCCCACCACCAAGAGGAGGAGGCCCATGTCATGAC 308  
QY 201 CTTCTGCTCTGACGGGTGTCTGTACAC-GTTCTTCTGAAAGCTTTTAGACAGTAACAGTTT 259  
Db 309 CTTCTGCTCTGACGGGTGTCTGTACACAGTTTCTTCTGAAAGCTTTTAGACAGTAACAGTTT 368  
QY 260 GCACATGAAAAATAAGCGAGCACCTAGACAA 291  
Db 369 GCACATGAAAAATAAGCGAGCACCTAGACAA 400  
RESULT 11  
ACH29871  
ID ACH29871 standard; cDNA; 485 BP.  
XX  
XX AC ACH29871;  
XX  
DT 13-OCT-2003 (first entry)  
XX  
XX DE Human testis cDNA #257.  
XX  
XX Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;

genome mapping; biodiversity; genetic disorder.  
XX  
XX Homo sapiens.  
XX  
XX US2003073623-A1.  
XX  
XX 17-APR-2003.  
XX  
XX 30-JUL-2001; 2001US-00918995.  
XX  
XX 30-JUL-2001; 2001US-00918995.  
XX  
XX (DRMA/) DRMANAC R T.  
XX (LABA/) LABAT I.  
XX (STAC/) STACHE-CRAIN B.  
XX (DICK/) DICKSON M C.  
XX (JONE/) JONES L W.  
XX  
XX Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;  
XX  
XX WPI; 2003-615964/58.  
XX  
XX New polynucleotide sequences obtained from various cDNA libraries, useful  
PT as hybridization probes, as oligomers for PCR, for chromosome and gene  
PT mapping, in the recombinant production of protein, or in generating  
PT antisense DNA or RNA.  
XX  
XX Claim 1; SEQ ID NO 17083; 44pp; English.  
XX  
XX The invention relates to an isolated polynucleotide comprising any one of  
CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was  
CC determined by the technique of SBH (sequencing by hybridisation). Also  
CC included is a purified polypeptide comprising a sequence corresponding to  
CC a reading frame of the novel polynucleotide. The nucleic acid sequences  
CC are useful in diagnostics as expressed sequence tags (EST) for  
CC identifying expressed genes or for physical mapping of the human genome,  
CC in forensics, in assessing biodiversity, or in identifying mutations  
CC responsible for genetic disorders and other traits. The nucleotide  
CC sequences are also useful as hybridisation probes, as oligomers for PCR,  
CC for chromosome and gene mapping, in the recombinant production of  
CC protein, or in generating antisense DNA or RNA. The purified polypeptide  
CC is useful for generating antibodies specific for it. The present sequence  
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data  
CC for this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from USPTO at  
CC seqdata.uspto.gov/sequence.html?DocID=20030073623  
XX  
XX SQ Sequence 485 BP; 122 A; 147 C; 122 G; 91 T; 0 U; 3 Other;  
Query Match 83.1%; Score 244.4; DB 9; Length 485;  
Best Local Similarity 99.6%; Pred. No. 2.6e-71;  
Matches 245; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 46 GGACAATCACCTACTGAGGAGCCCAACCCCGCCTCTGGATGGACCCGAATCAGGACTAG 105  
Db 75 GGACAATCACCTACTGAGGAGCCCAACCCCGCCTCTGGATGGACCCGAATCAGGACTAG 134  
QY 106 AGTTTAATCTGACGGGACATGTCATTCTCTATTCTTCTGTCACGGCCGCCCATGCCCA 165  
Db 135 AGTTTAATCTGACGGGACATGTCATTCTCTATTCTTCTGTCACGGCCGCCCATGCCCA 194  
QY 166 CCCCAACCAAGAGGAGGAGGCCCATGTCATGACCCCTCGTCTCTGACGGGTGTCTGTAC 225  
Db 195 CCCCAACCAAGAGGAGGAGGCCCATGTCATGACCCCTCGTCTCTGACGGGTGTCTGTAC 254  
QY 226 ACGTTCTTTTGAAAGCTTTTAGACAGTAACAGTTTGGACATGAAAAATAAGGAGCACCT 285  
Db 255 ACGTTCTTTTGAAAGCTTTTAGACAGTAACAGTTTGGACATGAAAAATAAGGAGCACCT 314  
QY 286 AGACAA 291  
Db 315 AAACAA 320



## RESULT 12

ADQ51104  
ID ADQ51104 standard; DNA; 466 BP.

AC ADQ51104;

DT 21-OCT-2004 (first entry)

XX Novel canine microarray-related DNA sequence SeqID2406.

XX canine microarray; drug screening; toxicity assay;  
XX environmental pollutant; cellular response; gene expression profile;  
XX toxic response; liver necrosis; fatty liver disease;  
XX protein adduct formation; hepatitis; dog; ds.

XX Canis familiaris.

XX WO2004063324-A2.

XX 29-JUL-2004.

XX 05-MAY-2003; 2003WO-US013853.

XX 03-MAY-2002; 2002US-0377240P.

XX (GENE-) GENE LOGIC INC.

XX (PFIZ ) PFIZER PROD INC.

XX Diggins JC, Porter M, Wei T;

XX WPI; 2004-561890/54.

XX New isolated nucleic acid molecule, useful for drug screening and  
XX toxicity assays or for assessing the impact, including toxicity, of a  
XX compound, pharmaceutical agent or environmental pollutant on a cell or  
XX living organism.

XX Claim 1; SEQ ID NO 2406; 41pp; English.

XX This invention is related to a novel isolated canine nucleic acid  
XX sequences and the construction of canine microarrays containing a  
XX significant portion of the canine genome. The isolated canine nucleic  
XX acid sequences of the invention may be useful for drug screening and  
XX toxicity assays. The invention is therefore useful for assessing the  
XX impact, including toxicity, of a compound, pharmaceutical agent or  
XX environmental pollutant on a cell or living organism. The methods are  
XX useful for detecting genes that are up- or down-regulated in canines in a  
XX disease state. The sequences are useful as diagnostic agents or markers  
XX to detect a cellular response in a sample individually or as part of a  
XX gene expression profile. It is also useful as a target for agents that  
XX modulate gene expression or activity. The database is useful for  
XX producing electronic Northern blots that allow the user to determine the cell  
XX type or tissue in which a given gene is expressed and to allow  
XX determination of the abundance or expression level of a given gene in a  
XX particular tissue or cell. The methods are useful for determining the  
XX similarity of a toxic response to one or more individual compounds. The  
XX methods are useful for predicting at least one toxic response or the  
XX likelihood that a compound or test agent will induce various specific  
XX pathologies such as those of the liver (liver necrosis, fatty liver  
XX disease, protein adduct formation or hepatitis), those of the kidney,  
XX heart, brain or testes, or other pathologies associated with at least one  
XX of the toxins. The methods are also useful for predicting or elucidating  
XX the potential cellular pathways influenced, induced or modulated by the  
XX compound or test agent due to the similarity of the expression profile  
XX compared to the profile induced by a known toxin. The present sequence is  
XX that of a canine DNA sequence which was claimed for use during the  
XX production of a canine microarray of the invention.

XX Sequence 466 BP; 102 A; 143 C; 113 G; 100 T; 0 U; 8 Other;

Query Match

Best Local Similarity 34.7%; Score 102; DB 13; Length 466;

Pred. No. 1.6e-23;

	Matches	182;	Conservative	0;	Mismatches	44;	Indels	16;	Gaps	4;
QY	21	TCCCATGATGCTCTCTCAAGGAGTGGACATACCTACTGAGAGCCCCACCCCGCCT	80							
DB	230	TCCCATGATGCTCTCTCAAGGAGTGGACATACCTACTGAGAGCCCCACCCCGCCT	284							
QY	81	CTGGATGGACCGAATCAGGAGCTAGAGTTTAACTGAGCGGAACTGTCTATTCTC---	136							
DB	285	TTGGACAGACCGAATCAGGAGCTAGAGTTTAACTGAGCGGAACTGTCTATTCTC---	344							
QY	137	TATTCTGTGACGCGCCCGCCATGCGCCCGCCACCCAGAGAGGAGGCGCCAGTCA	196							
DB	345	TATTCTGTGATGCGCCCGCCATGCGCCCGCCACCCAGAGAGGAGGCGCCAGTCA	404							
QY	197	TGACC-CTGCTCTCTGAGGGTG-----TCTGTACAGTTCTTCTGAAGAGCTTTAGACA	249							
DB	405	TGCCCTCATCTCTCTGAGGGTGTTCTCTATCTGTACACGCTCTTGTGAAGAGCTTTAGACA	464							
QY	250	GT 251								
DB	465	GT 466								

## RESULT 13

ABK13076/c

ID ABK13076 standard; DNA; 35973 BP.

XX AC ABK13076;

XX 23-APR-2002 (first entry)

XX Human amyloid beta precursor protein binding protein 1 gene.

XX Human; amyloid beta precursor protein binding protein 1; APPBP1; gene;  
XX Alzheimer's disease; transgenic animal; platelet aggregation;  
XX single nucleotide polymorphism; SNP; ds.

XX Homo sapiens.

Key	Location/Qualifiers
variation	replace(3436,G)
	/*tag= a
variation	/standard name= "Single nucleotide polymorphism"
	replace(3474,C)
	/*tag= b
CDS	/standard name= "Single nucleotide polymorphism"
	4085..31975
	/*tag= c
	/product= "APPBP1"
variation	/note= "Amyloid beta precursor protein binding protein 1"
	replace(11314,G)
	/*tag= d
variation	/standard name= "Single nucleotide polymorphism"
	replace(18123,G)
	/*tag= e
variation	/standard name= "Single nucleotide polymorphism"
	replace(18571,A)
	/*tag= f
variation	/standard name= "Single nucleotide polymorphism"
	replace(21229,A)
	/*tag= g
variation	/standard name= "Single nucleotide polymorphism"
	replace(21312,A)
	/*tag= h
variation	/standard name= "Single nucleotide polymorphism"
	replace(24300,A)
	/*tag= i
variation	/standard name= "Single nucleotide polymorphism"
	replace(26093,A)
	/*tag= j
variation	/standard name= "Single nucleotide polymorphism"
	replace(29334,C)
	/*tag= k



```
FT /standard_name= "Single nucleotide polymorphism"
FT replace(31499,T)
FT /*tag= 1
FT /standard_name= "Single nucleotide polymorphism"
FT replace(31582,G)
FT /*tag= m
FT /standard_name= "Single nucleotide polymorphism"
PN WO200202820-A1.
PN 10-JAN-2002.
XX
XX 02-JUL-2001; 2001WO-US020951.
XX
XX 30-JUN-2000; 2000US-0215511P.
XX
XX (GENA-) GENAISSANCE PHARM INC.
XX
XX Anastasio AE, Chew A, Choi JY, Kazemi A, Koshy B, Sausker EA;
XX Stephens CJ;
XX
XX WPI; 2002-164539/21.
XX P-PSDB; AAU11030.
XX
XX Amyloid beta precursor protein binding protein 159 kD (APPBP1) gene
XX polymorphic variants, useful e.g. in studying the expression and function
XX of APPBP1 and screening candidate drugs for treating Alzheimer's disease.
XX
XX Claim 1; Fig 1; 104pp; English.
XX
XX The invention relates to an isolated polypeptide comprising a sequence
XX which is a polymorphic variant of a reference sequence for the amyloid
XX beta precursor protein binding protein 1, 59kD (APPBP1) protein or its
XX fragment. The polymorphic variants are useful in studying the expression
XX and function of APPBP1, in expressing APPBP1 protein for use in screening
XX for candidate drugs to treat diseases related to APPBP1 activity, in
XX studying the effect of the variation on the biological activity of
XX APPBP1, and the binding affinity of candidate drugs targeting APPBP1 for
XX the treatment of disorders such as Alzheimer's disease. The haplotyping
XX methods are useful in validating APPBP1 as a candidate target for
XX treating a specific condition or disease predicted to be associated with
XX APPBP1 activity, or in the design of clinical trials of candidate drugs
XX for treating a specific condition or disease associated with APPBP1
XX activity. The transgenic animals are useful for studying expression of
XX the APPBP1 isogenes in vivo, for in vivo screening and testing of drugs
XX targeted against APPBP1 protein, and for testing the efficacy of
XX therapeutic agents and compounds for disorders related to platelet
XX aggregation in a biological system. The present sequence represents the
XX human APPBP1 gene as described in the method of the invention
XX
XX Sequence 35973 BP; 9909 A; 7327 C; 7959 G; 10766 T; 0 U; 12 Other;
XX
XX Query Match 12.2%; Score 35.8; DB 6; Length 35973;
XX Best Local Similarity 52.3%; Pred. No. 1.8;
XX Matches 79; Conservative 0; Mismatches 72; Indels 0; Gaps 0;
XX
XX QY 65 AGCCCCCCCCGCTCTGGTGGACCAATCAGGACTAGAGTTTAACTCAGCGGAAC 124
XX |||||
XX Db 7063 AGCAATTTCTCGCCCTCTAGAGTTGCTGGGACTACAGGCACAGCTGTACC 7004
XX |||||
XX QY 125 ATGTCATTTCTATTTCTGTGACGCGCCGCCATCCGCCACCCACCAAGAGGACGG 184
XX |||||
XX Db 7003 TGGCTTCATTTGTTTGTGCTGAGCATCATCTTTCCACTGATGCTCCATA 6944
XX |||||
XX QY 185 AGGGCCCAAGTCATGACCCCTCGTCTCTGCAGG 215
XX |||||
XX Db 6943 AGGGCCCAAGACAGCTGTCAAATGTTATGAGG 6913
XX |||||
XX
XX RESULT 14
XX ADR31515
XX ID ADR31515 standard; cDNA; 1559 BP.
XX
```

```
AC ADR31515;
XX
XX 02-DEC-2004 (first entry)
XX
XX Human adenosine deaminase cDNA.
XX
XX Human; adenosine deaminase; ADA; dideoxyinosine; ddi; AIDS; AZT;
XX 3'-azido-2',3'-dideoxythymidine; dideoxyadenosine deaminase; dda; ss;
XX gene.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 122..1213
XX FT /*tag= a
XX FT /product= "Adenosine deaminase"
XX
XX PN US2004175804-A1.
XX
XX 09-SEP-2004.
XX
XX 26-FEB-2004; 2004US-00787284.
XX
XX 04-MAR-2003; 2003US-0451842P.
XX
XX (SKON/) SKONEZNY P M.
XX (POLI/) POLITINO M.
XX (LIUS/) LIU S W
XX (BOYL/) BOYLE A W.
XX (CHEN/) CHEN J G.
XX (STEI/) STEIN G L.
XX (FRAN/) FRANCESCHINI T.
XX (ANDE/) ANDERSON W L.
XX
XX Skonezny PM, Politino M, Liu SW, Boyle AW, Chen JG, Stein GL;
XX Franceschini T, Anderson WL;
XX WPI; 2004-652027/63.
XX P-PSDB; ADR31514.
XX
XX Making dideoxyinosine (ddi), by obtaining enzyme expressing
XX dideoxyadenosine (dda) deaminase activity, immobilizing enzyme onto
XX insoluble support, contacting enzyme with dda solution to produce ddi
XX solution, isolating ddi.
XX
XX Claim 9; SEQ ID NO 2; 13pp; English.
XX
XX The invention relates to making 2',3'-dideoxyinosine (ddi), involves
XX obtaining an enzyme expressing dideoxyadenosine (dda) deaminase activity,
XX immobilizing the enzyme onto an insoluble support, contacting the enzyme
XX with a dda solution of at least 1% weight volume dda in water for a time
XX and under conditions to produce a ddi solution, and isolating the ddi
XX from the ddi solution. The enzyme is human adenosine deaminase (ADA)
XX appearing as ADR31514 and is encoded by the cDNA appearing as ADR31515 or
XX the E. coli-codon optimised cDNA appearing as ADR31516. The method is
XX useful for making 2',3'-dideoxyinosine (ddi) and provides a yield of at
XX least about 96% ddi that is at least about 99% pure. The ddi mother
XX liquor can be reused in subsequent runs. The human ADA is more stable
XX over a product pH range. ddi is used with AZT (3'-azido-2',3'-
XX dideoxythymidine) to treat AIDS. The present sequence is the wild-type
XX cDNA encoding human ADA.
XX
XX Sequence 1559 BP; 399 A; 448 C; 433 G; 279 T; 0 U; 0 Other;
XX
XX Query Match 11.0%; Score 32.4; DB 13; Length 1559;
XX Best Local Similarity 47.1%; Pred. No. 6.4;
XX Matches 99; Conservative 0; Mismatches 111; Indels 0; Gaps 0;
XX
XX QY 85 ATGGACCGAATCAGGACTAGAGTTTAACTCAGCGGAACATGTCATTTCTATTCTG 144
XX |||||
XX Db 1344 ATGTGTCATTTCTGCACACACACATACCTCGGATGCGCGGTCATCTCTGATTATG 1403
XX |||||
XX QY 145 TGACGCGCCCCCATGCCCCCACCACCCACCAAGAGCGAGGGGCCCATGATGACCCCTC 204
XX |||||
```

Db	1404	TGCCCTGGCCAGGACCGAGCGCCCTTCGCACATGGCATGGTTGAATCTGAACCCCTCCTT	1463
Qy	205	GTCTCTCAGGGTGTCTGTACACGCTTCTTCTGAAGCTTTAGACAGTAACAGTTTGACACA	264
Db	1464	CTGTGGCAACTGTGTACTGAAAAATCTGTGTCTCAATAAAGAAGCCCATGGCTGGTGGCATG	1523
Qy	265	TGAAAAATAAAGCGAGCACCTAGACAAAAA	294
Db	1524	CAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	1553

RESULT 15  
ABZ12539/c  
ID ABZ12539 standard; DNA; 2724 BP.  
XX  
AC ABZ12539;  
XX  
DT 21-JAN-2003 (first entry)  
XX  
DE Arabidopsis thaliana stress regulated gene SEQ ID NO 344.  
XX  
KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.  
XX  
OS Arabidopsis thaliana.  
XX  
PN WO200216655-A2.  
XX  
PD 28-FEB-2002.  
XX  
PP 24-AUG-2001; 2001WO-US026685.  
XX  
PR 24-AUG-2000; 2000US-0227866P.  
PR 26-JAN-2001; 2001US-0264647P.  
PR 22-JUN-2001; 2001US-0300111P.  
XX  
PA (SCRI ) SCRIPPS RES INST.  
PA (SYGN ) SYNGENTA PARTICIPATIONS AG.  
XX  
PI Harper JF, Kreps J, Wang X, Zhu T;  
XX WPI; 2002-304127/34.  
XX  
PT Identifying a stress condition to which a plant cell has been exposed and  
PT producing plants with increased tolerance to these abiotic stresses.  
XX  
PS Claim 144; SEQ ID NO 344; 577pp + Sequence Listing; English.  
XX  
CC The invention relates to identifying a stress condition to which a plant  
CC cell has been exposed, comprising: (a) contacting nucleic acid  
CC representative of expressed polynucleotides in the plant cell with an  
CC array or probes representative of the plant cell genome; and (b)  
CC detecting a profile of expressed polynucleotides in the plant cell  
CC characteristic of a stress response. The method is useful in the  
CC production of transgenic plants, cells and seeds and in producing plants  
CC with increased tolerance to abiotic stress. The present sequence is that  
CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used  
CC in methods of the invention. Note: The sequence data for this patent is  
CC not represented in the printed specification but is based on sequence  
CC information supplied to Derwent by the European Patent Office  
XX  
SQ Sequence 2724 BP; 751 A; 558 C; 676 G; 739 T; 0 U; 0 Other;

Query Match	11.0%;	Score 32.4;	DB 6;	Length 2724;
Best Local Similarity	60.0%;	Pred. No. 8.1;		
Matches	54;	Conservative 0;	Mismatches 36;	Indels 0; Gaps 0;
Qy	192	AGTCATGACCTCGCTCTCGAGGTGCTGTACACGTTCTTCTGAAGCTTTAGACAGT	251	
Db	2487	AGGAATGATCCCAACACGAGGTGCTCGATGATTCTCTCAAAGCTCTTGGCAAT	2428	
Qy	252	AACAGTTTGACATGAAAAATAAAGCGAGC	281	

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 10:04:41 ; Search time 140.979 Seconds  
(without alignments)  
4072.612 Million cell updates/sec

Title: US-09-980-046B-6  
Perfect score: 323  
Sequence: 1 cggggccaggctatgcct.....gtgcacgtgttgtaaaaa 323

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA.\*  
1: /cgn2\_6/ptodata/1/ina/1 COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5 COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/H COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/PP COMB.seq.\*  
7: /cgn2\_6/ptodata/1/ina/RE COMB.seq.\*  
8: /cgn2\_6/ptodata/1/ina/RE COMB.seq.\*  
9: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	311.8	96.5	764	3	US-08-991-789A-283
C 2	311.8	96.5	764	3	US-09-062-451-283
C 3	311.8	96.5	764	3	US-09-289-198-283
C 4	311.8	96.5	764	3	US-09-429-755-283
C 5	311.8	96.5	764	3	US-09-699-295-283
C 6	311.8	96.5	764	3	US-09-534-825A-283
C 7	85	26.3	642	3	US-09-370-838-119
C 8	85	26.3	642	3	US-09-854-133-119
C 9	84.8	26.3	808	3	US-09-023-655-1279
C 10	84.8	26.3	808	3	US-09-461-912A-25
C 11	76	23.5	362	3	US-09-461-912A-2
C 12	52.6	16.3	362	3	US-09-461-912A-1
C 13	50	15.5	50	3	US-10-131-827-2535
C 14	34.6	10.7	289	3	US-09-007-005-17
C 15	34.6	10.7	289	3	US-09-244-796-17
C 16	34.6	10.7	399	3	US-09-621-976-8976
C 17	34.6	10.7	2441	3	US-09-949-016-3807
C 18	34.2	10.6	23155	3	US-09-949-016-12366
C 19	34	10.5	25733	3	US-09-902-540-1215
C 20	32.6	10.1	152481	3	US-09-949-016-12521
C 21	32.6	10.1	152798	3	US-09-949-016-12775
C 22	32.6	10.1	152822	3	US-09-949-016-17518
C 23	32.6	10.1	152822	3	US-09-949-016-17519
C 24	32.4	10.0	3306	3	US-09-081-385-10

25	32.4	10.0	3306	3	US-09-752-639-10	Sequence 10, Appl
26	32.4	10.0	3306	3	US-09-712-813-10	Sequence 10, Appl
27	32.4	10.0	3306	3	US-09-700-354A-10	Sequence 10, Appl
C 28	31.6	9.8	9279	3	US-09-487-558B-171	Sequence 171, App
29	31.2	9.7	847	3	US-09-533-559-2149	Sequence 2149, App
30	31	9.6	1041	3	US-09-248-796A-1489	Sequence 1489, App
31	31	9.6	1230	3	US-09-252-991A-947	Sequence 947, App
32	31	9.6	1521	3	US-09-252-991A-989	Sequence 989, App
C 33	31	9.6	2201	3	US-09-149-476-245	Sequence 245, App
C 34	30.8	9.5	660	3	US-09-902-540-3995	Sequence 3995, App
35	30.8	9.5	1023	3	US-09-543-681A-1606	Sequence 1606, App
36	30.8	9.5	7218	2	US-08-232-463-14	Sequence 14, Appl
37	30.8	9.5	19862	3	US-09-902-540-1198	Sequence 1198, App
C 38	30.6	9.5	534	3	US-09-593-995-6	Sequence 6, Appl
C 39	30.6	9.5	534	3	US-10-139-667-6	Sequence 6, Appl
40	30.6	9.5	1068	3	US-09-170-496D-129	Sequence 129, App
41	30.6	9.5	1068	3	US-09-170-496D-231	Sequence 231, App
42	30.6	9.5	1161	2	US-08-153-848-31	Sequence 31, Appl
43	30.6	9.5	1161	6	PCT-US93-11153-31	Sequence 31, Appl
44	30.6	9.5	1320	2	US-08-599-252-84	Sequence 84, Appl
45	30.6	9.5	1320	2	US-08-436-074-57	Sequence 57, Appl

ALIGNMENTS

RESULT 1  
US-08-991-789A-283/c  
; Sequence 283, Application US/08991789A  
; Patent No/ 6225054

GENERAL INFORMATION:  
APPLICANT: Prudakis, Tony N.  
Smith, John M.  
Reed, Steven G.

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE

TREATMENT AND DIAGNOSIS OF BREAST CANCER

NUMBER OF SEQUENCES: 292

CORRESPONDENCE ADDRESS:

ADDRESSEE: Seed IP Law Group

STREET: 701 Fifth Avenue, Suite 6300

CITY: Seattle

STATE: Washington

COUNTRY: USA

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/991,789A

FILING DATE: 11-Dec-1997

CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Potter, Jane E. R.

REGISTRATION NUMBER: 33,332

REFERENCE/DOCKET NUMBER: 210121.419C3

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 283:

SEQUENCE CHARACTERISTICS:

LENGTH: 764 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 283:

US-08-991-789A-283

Query Match 96.5%; Score 311.8; DB 3; Length 764;

Best Local Similarity 94.7%; Pred. No. 2.6e-93;

Matches 306; Conservative 14; Mismatches 3; Indels 0; Gaps 0;

Qy	1	CGGGGCCAGGCGTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCGGG	60
Db	330	CGGGCMCAGGCGTATGCCCCACCGCAAGTGCCTGAASATSTGGGCCCTGATTSTGGG	271
Qy	61	CATCTCATGACGATTGGATTCCACCTGTTACTGGTATTTCGGCTCTGTGACAGTCTACCA	120
Db	270	CATCTCTRTGACGCTTTGGATTTCATGCTGTDAATGGWATTCGGCTCTGTGACAGTCTACCA	211
Qy	121	TATTATGTTACAGATAATAACAGAAAAACGGGGTTACTAGTAGCGGCCCATAGCCTGCAA	180
Db	210	TATTATGTTACAGATAATAACAGAAAAACGGGGTTABTAGTAGCGGCCCATAGCCTGCAA	151
Qy	181	CTTTGCACTCCACTGTGCAATGCTGGCCCTGACGCTGGGGCTGTGGCCCTGCCCCCT	240
Db	150	CTTTGCACTCCACTGTGCAATGCTGGCCCTGACGCTGGGGCTGTGGCCCTGCCCCCT	91
Qy	241	TGCTCTCGCCCTTAGATACAGCAGTTTATACCACACACCTGTCTACAGTGTCAATCAAT	300
Db	90	TGCTCTCGCCCTTAGATACAGCAGTTTATACCACACACCTGTCTACAGTGTCAATCAAT	31
Qy	301	AAAGTGCACGTGTGTGAAAAA	323
Db	30	AAAGTGCACGTGTGTGAAAAA	8

Qy 121 TATTATGTTACAGATAAATACAGSAAAAACGGGGTTACTAGTAGCGGCCCATAGCCTGCAA 180  
Db |||||  
Qy 210 TATTATGTTACAGATAAATACAGSAAAAACGGGGTTACTAGTAGCGGCCCATAGCCTGCAA 151  
Db |||||  
Qy 181 CTTTGGCACTCAGTGTCAATGTGCGCCCTGCAGCTGGGGCTGTGGCCCTGCCCT 240  
Db |||||  
Qy 150 CTTTGGCACTCAGTGTCAATGTGCGCCCTGCAGCTGGGGCTGTGGCCCTGCCCT 91  
Db |||||  
Qy 241 TGGTCCTGCCCTAGATACAGAGTTTATATACCCACACACCTGTCTACAGTGTCAATCAAT 300  
Db |||||  
Qy 90 TGGTCCTGCCCTAGATACAGAGTTTATATACCCACACACCTGTCTACAGTGTCAATCAAT 31  
Db |||||  
Qy 301 AAAGTGCACGTGCTTGTGAAAAA 323  
Db |||||  
Qy 30 AAAGTGCACGTGCTTGTGAAAAA 8  
Db |||||

## RESULT 4

US-09-429-755-283/c  
; Sequence 283, Application US/09429755A  
; Patent No. 6656480

## GENERAL INFORMATION:

; APPLICANT: Frudakis, Tony N.  
; APPLICANT: Smith, John M.  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Mishner, Lynda  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Dillon, Davin C.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE

; FILE REFERENCE: 210121.419C6

; CURRENT APPLICATION NUMBER: US/09/429,755A

; CURRENT FILING DATE: 1999-10-28

; NUMBER OF SEQ ID NOS: 315

; SOFTWARE: FastSEQ for Windows Version 3.0

; SEQ ID NO 283

; LENGTH: 764

; TYPE: DNA

; ORGANISM: Homo sapien

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (1)...(764)

; OTHER INFORMATION: n = A,T,C or G

## US-09-429-755-283

Query Match 96.5%; Score 311.8; DB 3; Length 764;  
Best Local Similarity 94.7%; Pred. No. 2.6e-93;  
Matches 306; Conservative 14; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CGGGGCCAGGCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTTCTGGG 60  
Db |||||  
Qy 330 CGGGGCMAGGCTATGCCCCCAGCCGCAAGTGCCTGAASATSTGGGCCCTGATTTSTGGG 271  
Db |||||  
Qy 61 CATCCTCATGACCAATGGATTCAACCTGTACTGTGTTTACGCTCTGTGACAGTCTACCA 120  
Db |||||  
Qy 270 CATCCTCTGACCTTGGATTTCATCTGTDAYTGGWATTGGCTCTGTGACAGTCTACCA 211  
Db |||||  
Qy 121 TATTATGTTACAGATAAATACAGSAAAAACGGGGTTACTAGTAGCGGCCCATAGCCTGCAA 180  
Db |||||  
Qy 210 TATTATGTTACAGATAAATACAGSAAAAACGGGGTTACTAGTAGCGGCCCATAGCCTGCAA 151  
Db |||||  
Qy 181 CTTTGGCACTCAGTGTCAATGTGCGCCCTGCAGCTGGGGCTGTGGCCCTGCCCT 240  
Db |||||  
Qy 150 CTTTGGCACTCAGTGTCAATGTGCGCCCTGCAGCTGGGGCTGTGGCCCTGCCCT 91  
Db |||||  
Qy 241 TGGTCCTGCCCTAGATACAGAGTTTATATACCCACACACCTGTCTACAGTGTCAATCAAT 300  
Db |||||  
Qy 90 TGGTCCTGCCCTAGATACAGAGTTTATATACCCACACACCTGTCTACAGTGTCAATCAAT 31  
Db |||||  
Qy 301 AAAGTGCACGTGCTTGTGAAAAA 323  
Db |||||  
Qy 30 AAAGTGCACGTGCTTGTGAAAAA 8  
Db |||||

## RESULT 5

US-09-699-295-283/c  
; Sequence 283, Application US/09699295  
; Patent No. 6828431

## GENERAL INFORMATION:

; APPLICANT: Frudakis, Tony N.  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Smith, John M.  
; APPLICANT: Mishner, Lynda  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Harlocker, Susan L.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE

; FILE REFERENCE: 210121.419C10

; CURRENT APPLICATION NUMBER: US/09/699,295

; CURRENT FILING DATE: 2000-10-26

; NUMBER OF SEQ ID NOS: 326

; SOFTWARE: FastSEQ for Windows Version 3.0

; SEQ ID NO 283

; LENGTH: 764

; TYPE: DNA

; ORGANISM: Homo sapien

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (1)...(764)

; OTHER INFORMATION: n = A,T,C or G

## US-09-699-295-283

Query Match 96.5%; Score 311.8; DB 3; Length 764;  
Best Local Similarity 94.7%; Pred. No. 2.6e-93;  
Matches 306; Conservative 14; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CGGGGCCAGGCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTTCTGGG 60  
Db |||||  
Qy 330 CGGGGCMAGGCTATGCCCCCAGCCGCAAGTGCCTGAASATSTGGGCCCTGATTTSTGGG 271  
Db |||||  
Qy 61 CATCCTCATGACCAATGGATTCAACCTGTACTGTGTTTACGCTCTGTGACAGTCTACCA 120  
Db |||||  
Qy 270 CATCCTCTGACCTTGGATTTCATCTGTDAYTGGWATTGGCTCTGTGACAGTCTACCA 211  
Db |||||  
Qy 121 TATTATGTTACAGATAAATACAGSAAAAACGGGGTTACTAGTAGCGGCCCATAGCCTGCAA 180  
Db |||||  
Qy 210 TATTATGTTACAGATAAATACAGSAAAAACGGGGTTACTAGTAGCGGCCCATAGCCTGCAA 151  
Db |||||  
Qy 181 CTTTGGCACTCAGTGTCAATGTGCGCCCTGCAGCTGGGGCTGTGGCCCTGCCCT 240  
Db |||||  
Qy 150 CTTTGGCACTCAGTGTCAATGTGCGCCCTGCAGCTGGGGCTGTGGCCCTGCCCT 91  
Db |||||  
Qy 241 TGGTCCTGCCCTAGATACAGAGTTTATATACCCACACACCTGTCTACAGTGTCAATCAAT 300  
Db |||||  
Qy 90 TGGTCCTGCCCTAGATACAGAGTTTATATACCCACACACCTGTCTACAGTGTCAATCAAT 31  
Db |||||  
Qy 301 AAAGTGCACGTGCTTGTGAAAAA 323  
Db |||||  
Qy 30 AAAGTGCACGTGCTTGTGAAAAA 8  
Db |||||

## RESULT 6

US-09-534-825A-283/c  
; Sequence 283, Application US/09534825A  
; Patent No. 6861506

## GENERAL INFORMATION:

; APPLICANT: Frudakis, Tony N.  
; APPLICANT: Smith, John M.  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Mishner, Lynda  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Dillon, Davin C.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE

1 CCGGGCCCGCCCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCCTGGG 60  
2 FILE REFERENCE: 210121.419C7  
3 CURRENT APPLICATION NUMBER: US/09/534,825A  
4 CURRENT FILING DATE: 2000-03-23  
5 NUMBER OF SEQ ID NOS: 317  
6 SOFTWARE: FastSeq for Windows Version 3.0  
7 SEQ ID NO 283  
8 LENGTH: 764  
9 TYPE: DNA  
10 ORGANISM: Homo sapien  
11 NAME/KEY: misc feature  
12 LOCATION: (1)-(764)  
13 OTHER INFORMATION: n = A,T,C or G  
14 US-09-534-825A-283

Query Match 96.5%; Score 311.8; DB 3; Length 764;  
Best Local Similarity 94.7%; Pred. No. 2.6e-93;  
Matches 306; Conservative 14; Mismatches 3; Indels 0; Gaps 0;

1 CCGGGCCCGCCCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCCTGGG 60  
2 Db 330 CCGGGCCCGCCCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCCTGGG 271  
3 Qy 61 CATCCTCATGACCAATTGGATTACCCCTGTTACTGCTATTGCGCTCTGTGACAGTCTACCA 120  
4 Db 270 CATCCTCATGACCAATTGGATTACCCCTGTTACTGCTATTGCGCTCTGTGACAGTCTACCA 211  
5 Qy 121 TATTATGTTACAGATAATACAGGAAACCGGGTTACTAGTAGCGGCCCATAGCCTGCAA 180  
6 Db 210 TATTATGTTACAGATAATACAGGAAACCGGGTTACTAGTAGCGGCCCATAGCCTGCAA 151  
7 Qy 181 CCTTTGCACCTCCACTGTGCAATGCTGGCCCTGACCGTGGGGCTGTGGCCCTGCCCCCT 240  
8 Db 150 CCTTTGCACCTCCACTGTGCAATGCTGGCCCTGACCGTGGGGCTGTGGCCCTGCCCCCT 91  
9 Qy 241 TGGTCTGCGCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 300  
10 Db 90 TGGTCTGCGCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 31  
11 Qy 301 AAAGTGCACGTGCTTGTGAAAAA 323  
12 Db 30 AAAGTGCACGTGCTTGTGAAAAA 8

RESULT 7  
US-09-370-838-119  
; Sequence 119, Application US/09370838  
; Patent No. 644425  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Mohamath, Raodoh  
; APPLICANT: Secrist, Heather  
; TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF  
; LUNG CANCER AND METHODS FOR THEIR USE  
; FILE REFERENCE: 210121.475C1  
; CURRENT APPLICATION NUMBER: US/09/370,838  
; CURRENT FILING DATE: 1999-08-09  
; EARLIER APPLICATION NUMBER: US 09/285,323  
; NUMBER OF SEQ ID NOS: 289  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 119  
; LENGTH: 642  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-370-838-119

Query Match 26.3%; Score 85; DB 3; Length 642;  
Best Local Similarity 60.7%; Pred. No. 5.7e-18;  
Matches 196; Conservative 0; Mismatches 90; Indels 37; Gaps 2;

1 CCGGGCCCGCCCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCCTGGG 60  
2 Db 332 CCGGGCCCGCCCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCCTGGG 391  
3 Qy 61 CATCCTCATGACCAATTGGATTACCCCTGTTACTGCTATTGCGCTCTGTGACAGTCTACCA 120  
4 Db 392 CATCCTCATGACCAATTGCTCTCATGCTCATCCAGTGTGATCTTCCAGGCTATGGATA 451  
5 Qy 121 TATTATGTTACAGATAATACAGGAAACCGGGTTACTAGTAGCGGCCCATAGCCTGCAA 180  
6 Db 452 GATCAGG-----AGGCATCAGTACAGGCCAGGAGCTCTGCCCATGACCTGTAT 498  
7 Qy 181 CCTTTGCACCTCCACTGTGCAATGCTGGCCCTGACCGTGGGGCTGTGGCCCTGCCCCCT 240  
8 Db 499 CCACGCTACTCCAACTTCCATTCCTGCCCCCTG-----CCCC 534  
9 Qy 241 TGGTCTGCGCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 300  
10 Db 535 CCGAGCGGAGTCTGTATCAGCCCTTTATCCTCACACGCTTTTCTACAATGGCATTCAAT 594  
11 Qy 301 AAAGTGCACGTGCTTGTGAAAAA 323  
12 Db 595 AAAGTGCACGTGCTTGTGAAAAA 617

RESULT 8  
US-09-854-133-119  
; Sequence 119, Application US/09854133  
; Patent No. 6759508  
; GENERAL INFORMATION:  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Mohamath, Raodoh  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Benson, Darin R.  
; APPLICANT: Secrist, Heather  
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.475C10  
; CURRENT APPLICATION NUMBER: US/09/854,133  
; NUMBER OF SEQ ID NOS: 735  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 119  
; LENGTH: 642  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-854-133-119

Query Match 26.3%; Score 85; DB 3; Length 642;  
Best Local Similarity 60.7%; Pred. No. 5.7e-18;  
Matches 196; Conservative 0; Mismatches 90; Indels 37; Gaps 2;  
1 CCGGGCCCGCCCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCCTGGG 60  
2 Db 332 CCGGGCCCGCCCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCCTGGG 391  
3 Qy 61 CATCCTCATGACCAATTGGATTACCCCTGTTACTGCTATTGCGCTCTGTGACAGTCTACCA 120  
4 Db 392 CATCCTCATGACCAATTGCTCTCATGCTCATCCAGTGTGATCTTCCAGGCTATGGATA 451  
5 Qy 121 TATTATGTTACAGATAATACAGGAAACCGGGTTACTAGTAGCGGCCCATAGCCTGCAA 180  
6 Db 452 GATCAGG-----AGGCATCAGTACAGGCCAGGAGCTCTGCCCATGACCTGTAT 498  
7 Qy 181 CCTTTGCACCTCCACTGTGCAATGCTGGCCCTGACCGTGGGGCTGTGGCCCTGCCCCCT 240  
8 Db 499 CCACGCTACTCCAACTTCCATTCCTGCCCCCTG-----CCCC 534  
9 Qy 241 TGGTCTGCGCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 300  
10 Db 535 CCGAGCGGAGTCTGTATCAGCCCTTTATCCTCACACGCTTTTCTACAATGGCATTCAAT 594  
11 Qy 301 AAAGTGCACGTGCTTGTGAAAAA 323

```
Db      595 AAGTGCACGTGTTCTGGTGA 617
|||||
RESULT 9
US-09-023-655-1279
; Sequence 1279, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Sellhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1279:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 808 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: G311374
US-09-023-655-1279

Query Match      26.3%; Score 84.8; DB 3; Length 808;
Best Local Similarity 61.0%; Pred. No. 7.4e-18;
Matches 194; Conservative 0; Mismatches 87; Indels 37; Gaps 2;

Qy      1 CGGGGCCAGCGCTATGCTCCACCGCAAGTGCCTGAAACATCTGGGCCCTGATTCGGG 60
Db      519 CGGGGCCAGCGCTATGCTCCACCGCAAGTGCCTGAAACATCTGGGCCCTGATTCGGG 578

Qy      61 CATCTCATGACCATTTGGATTACCCCTGTTACTGGTATTTCGGCTCTGTGACAGTCTACCA 120
Db      579 CATCTCATGACCATTTGGATTACCCCTGTTACTGGTATTTCGGCTCTGTGACAGTCTATGGATA 638

Qy      121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCCTGCAA 180
Db      639 GATCAGG-----AGGCATCACTGAGGCCAGGAGCTCTGCCCCATGACCTGTAT 685

Qy      181 CTTTGCACCTCCACTGTGCAATGCTGGCCCTGCAAGCTGGGGCTGTGGCCCTGCCCCCT 240
Db      686 CCCAGCTACTCCAACTTCCCATTTCTCGCCCTG-----CCCC 721

Qy      241 TGGTCTCGCCCTAGATACAGACAGTCTTATACCCACACACACCTGTCTACAGTGTCTATTCAAT 300
Db      722 CGGAGCGAGTCTCTGTATCAGCCCTTTATCTCTACACGCTTTTCTACAAATGGCATTCAAT 781

Qy      301 AAAGTGCACGTGTTGTG 318
Db      782 AAAGTGCACGTGTTTCTG 799

RESULT 10
US-09-461-912A-25
; Sequence 25, Application US/09461912A
; Patent No. 6709855
; GENERAL INFORMATION:
; APPLICANT: Stanton, Lawrence A.
; APPLICANT: White, R. Tyler
; APPLICANT: Damm, Deborah L.
; APPLICANT: Lewicki, John A.
; TITLE OF INVENTION: Methods for detection and use of
; TITLE OF INVENTION: differentially expressed genes in disease states
; FILE REFERENCE: SCIOS.011A
; CURRENT APPLICATION NUMBER: US/09/461,912A
; CURRENT FILING DATE: 1999-12-15
; PRIOR APPLICATION NUMBER: US 60/113,008
; PRIOR FILING DATE: 1998-12-18
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 808
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (238)....(639)
US-09-461-912A-25

Query Match      26.3%; Score 84.8; DB 3; Length 808;
Best Local Similarity 61.0%; Pred. No. 7.4e-18;
Matches 194; Conservative 0; Mismatches 87; Indels 37; Gaps 2;

Qy      1 CGGGGCCAGCGCTATGCTCCACCGCAAGTGCCTGAAACATCTGGGCCCTGATTCGGG 60
Db      519 CGGGGCCAGCGCTATGCTCCACCGCAAGTGCCTGAAACATCTGGGCCCTGATTCGGG 578

Qy      61 CATCTCATGACCATTTGGATTACCCCTGTTACTGGTATTTCGGCTCTGTGACAGTCTACCA 120
Db      579 CATCTCATGACCATTTGGATTACCCCTGTTACTGGTATTTCGGCTCTGTGACAGTCTATGGATA 638

Qy      121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCCTGCAA 180
Db      639 GATCAGG-----AGGCATCACTGAGGCCAGGAGCTCTGCCCCATGACCTGTAT 685

Qy      181 CTTTGCACCTCCACTGTGCAATGCTGGCCCTGCAAGCTGGGGCTGTGGCCCTGCCCCCT 240
Db      686 CCCAGCTACTCCAACTTCCCATTTCTCGCCCTG-----CCCC 721

Qy      241 TGGTCTCGCCCTAGATACAGACAGTCTTATACCCACACACACCTGTCTACAGTGTCTATTCAAT 300
Db      722 CGGAGCGAGTCTCTGTATCAGCCCTTTATCTCTACACGCTTTTCTACAAATGGCATTCAAT 781

Qy      301 AAAGTGCACGTGTTGTG 318
Db      782 AAAGTGCACGTGTTTCTG 799

RESULT 11
US-09-461-912A-2
; Sequence 2, Application US/09461912A
; Patent No. 6709855
; GENERAL INFORMATION:
; APPLICANT: Stanton, Lawrence A.
; APPLICANT: White, R. Tyler
; APPLICANT: Damm, Deborah L.
```

```
; APPLICANT: Lewicki, John A.
; TITLE OF INVENTION: Methods for detection and use of
; FILE REFERENCE: SCIOS.011A
; CURRENT APPLICATION NUMBER: US/09/461.912A
; CURRENT FILING DATE: 1999-12-15
; PRIOR FILING DATE: 1999-12-18
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 362
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-461-912A-2

Query Match      23.5%; Score 76; DB 3; Length 362;
Best Local Similarity 100.0%; Pred. No. 4.1e-15;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGGCCCGAGCCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCGGG 60
Db 285 CGGGCCCGAGCCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCGGG 344

Qy 61 CATCTCATGACCATT 76
Db 345 CATCTCATGACCATT 360

RESULT 12
US-09-461-912A-1
; Sequence 1, Application US/09461912A
; Patent No. 6709855
; GENERAL INFORMATION:
; APPLICANT: Stanton, Lawrence A.
; APPLICANT: White, R. Tyler
; APPLICANT: Damm, Deborah L.
; APPLICANT: Lewicki, John A.
; TITLE OF INVENTION: Methods for detection and use of
; FILE REFERENCE: SCIOS.011A
; CURRENT APPLICATION NUMBER: US/09/461.912A
; CURRENT FILING DATE: 1999-12-15
; PRIOR FILING DATE: 1999-12-18
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 362
; TYPE: DNA
; ORGANISM: Rattus rattus
US-09-461-912A-1

Query Match      16.3%; Score 52.6; DB 3; Length 362;
Best Local Similarity 81.3%; Pred. No. 2.5e-07;
Matches 61; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 2 GGGGCCCGAGCCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCGGG 61
Db 286 GGGGCCCGAGCCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCGGG 345

Qy 62 ATCTCATGACCATT 76
Db 346 ATCTCATGACCATT 360

RESULT 13
US-10-131-827-2535
; Sequence 2535, Application US/10131827
; Patent No. 6905827
; GENERAL INFORMATION:
; APPLICANT: Wohlgenuth, Jay
; APPLICANT: Fry, Kirk
```

```
; APPLICANT: Woodward, Robert
; APPLICANT: Ly, Ngoc
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND MONITORING AUTOIMMUNE
; FILE REFERENCE: CHRONIC INFLAMMATORY DISEASES
; CURRENT APPLICATION NUMBER: US/10/131,827
; CURRENT FILING DATE: 2002-09-06
; PRIOR FILING DATE: 2002-09-06
; NUMBER OF SEQ ID NOS: 290
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2535
; LENGTH: 50
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-131-827-2535

Query Match      15.5%; Score 50; DB 3; Length 50;
Best Local Similarity 100.0%; Pred. No. 6.5e-07;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 250 CCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGCATTCAA 299
Db 1 CCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGCATTCAA 50

RESULT 14
US-09-007-005-17/c
; Sequence 17, Application US/09007005B
; Patent No. 6258558
; GENERAL INFORMATION:
; APPLICANT: Szostak, Jack W.
; APPLICANT: Roberts, Richard W.
; APPLICANT: Liu, Rihe
; TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
; FILE REFERENCE: 00786/350003
; CURRENT APPLICATION NUMBER: US/09/007,005B
; CURRENT FILING DATE: 1998-01-14
; EARLIER APPLICATION NUMBER: 60/035,963
; EARLIER FILING DATE: 1997-01-27
; EARLIER APPLICATION NUMBER: 60/064,491
; EARLIER FILING DATE: 1997-11-06
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 289
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Translation template
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(289)
; OTHER INFORMATION: n = A,T,C or G
US-09-007-005-17

Query Match      10.7%; Score 34.6; DB 3; Length 289;
Best Local Similarity 5.8%; Pred. No. 0.21;
Matches 10; Conservative 78; Mismatches 83; Indels 0; Gaps 0;

Qy 148 ACGGGGTACTAGTAGCCGCCATAGCCTGCACACCTTTGCACCTCCACATGCTGG 207
Db 250 AYAYGAYGYTTYAYCYGYCYAYGYCYTGYSYNYNYNYNYNYNYNYNYNYNY 191

Qy 208 CCTGCACGCTGGGGCTGTTGCCCTTGGTCCCTGCCCTAGATACAGCATTT 267
Db 190 NYSYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNY 131

Qy 268 ATACCCACACACCTGTCTACAGTGTCTATTCATTAATAAGTCACGCTGTGTG 318
```





***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 10:04:41 ; Search time 38.4091 Seconds  
(without alignments)  
4072.612 Million cell updates/sec

Title: US-09-980-046B-5  
Perfect score: 88  
Sequence: 1 cggagtcgagccctgtatca.....cgtgtttcttgggtacaaaaa 88

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

- Database : Issued Patents NA:\*
- 1: /cgn2\_6/ptodata/1/ina/1 COMB.seq:\*
  - 2: /cgn2\_6/ptodata/1/ina/5 COMB.seq:\*
  - 3: /cgn2\_6/ptodata/1/ina/6A COMB.seq:\*
  - 4: /cgn2\_6/ptodata/1/ina/6B COMB.seq:\*
  - 5: /cgn2\_6/ptodata/1/ina/H COMB.seq:\*
  - 6: /cgn2\_6/ptodata/1/ina/PTUS COMB.seq:\*
  - 7: /cgn2\_6/ptodata/1/ina/PP COMB.seq:\*
  - 8: /cgn2\_6/ptodata/1/ina/RE COMB.seq:\*
  - 9: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	80	90.9	642	US-09-370-838-119	Sequence 119, App
2	80	90.9	642	US-09-854-133-119	Sequence 119, App
3	75.8	86.1	808	US-09-023-655-1279	Sequence 1279, Ap
4	75.8	86.1	808	US-09-461-912A-25	Sequence 25, Appl
5	44.8	50.9	50	US-10-131-827-2927	Sequence 2927, Ap
6	42.8	48.6	764	US-08-991-789A-283	Sequence 283, App
7	42.8	48.6	764	US-09-062-451-283	Sequence 283, App
8	42.8	48.6	764	US-09-289-198-283	Sequence 289, App
9	42.8	48.6	764	US-09-429-755-283	Sequence 283, App
10	42.8	48.6	764	US-09-699-295-283	Sequence 283, App
11	42.8	48.6	764	US-09-534-825A-283	Sequence 283, App
12	27.4	31.1	25202	US-09-949-016-13151	Sequence 13151, A
13	26.8	30.5	513	US-09-702-705-314	Sequence 314, App
14	26.8	30.5	513	US-09-736-457-314	Sequence 314, App
15	26.8	30.5	513	US-09-614-124B-314	Sequence 314, App
16	26.8	30.5	513	US-09-671-325-314	Sequence 314, App
17	26.8	30.5	513	US-09-589-184-314	Sequence 314, App
18	26.8	30.5	513	US-09-658-824-314	Sequence 314, App
19	26.8	30.5	513	US-10-017-754-314	Sequence 314, App
20	26.8	30.5	513	US-09-651-563-314	Sequence 314, App
21	26.8	30.5	513	US-09-519-642-314	Sequence 314, App
22	26.8	30.5	601	US-09-949-016-148364	Sequence 148364, A
23	26.8	30.5	601	US-09-949-016-167312	Sequence 167312, A
24	26.8	30.5	40237	US-09-949-016-16461	Sequence 16461, A

25	26.8	30.5	108169	3	US-09-949-016-12898	Sequence 12898, A
26	26.8	30.5	108169	3	US-09-949-016-15907	Sequence 15907, A
27	26.6	30.2	50	3	US-10-131-827-2535	Sequence 2535, Ap
28	26.4	30.0	601	3	US-09-949-016-108834	Sequence 108834, A
29	26.4	30.0	601	3	US-09-949-016-108835	Sequence 108835, A
c 30	26.4	30.0	60137	3	US-09-949-016-14735	Sequence 14735, A
c 31	26.4	30.0	60137	3	US-09-949-016-14912	Sequence 14912, A
32	25.4	28.9	3382	3	US-09-513-505-5	Sequence 5, Appli
33	25.4	28.9	3382	3	US-09-513-505-7	Sequence 7, Appli
34	25.4	28.9	3420	3	US-09-513-505-1	Sequence 1, Appli
35	25.4	28.9	3478	3	US-09-513-505-3	Sequence 3, Appli
36	25.4	28.9	43360	3	US-09-453-702B-206	Sequence 206, App
37	25.4	28.9	43360	3	US-10-114-170-206	Sequence 206, App
38	25.4	28.9	45325	3	US-09-453-702B-261	Sequence 261, App
39	25.4	28.9	45325	3	US-10-114-170-261	Sequence 261, App
40	25.2	28.6	1757	3	US-09-484-970B-37	Sequence 37, Appl
c 41	25.2	28.6	101951	3	US-09-949-016-15648	Sequence 15648, A
c 42	25	28.4	601	3	US-09-949-016-171564	Sequence 171564, A
c 43	25	28.4	601	3	US-09-949-016-171565	Sequence 171565, A
c 44	25	28.4	691	3	US-09-949-016-2545	Sequence 2545, Ap
c 45	25	28.4	695	3	US-09-513-999C-3892	Sequence 3892, Ap

ALIGNMENTS

RESULT 1  
US-09-370-838-119  
; Sequence 119, Application US/09370838  
; Patent No. 6444425  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Mohamath, Roadoh  
; APPLICANT: Secrist, Heather  
; TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF  
; FILE REFERENCE: 210121.475C1  
; CURRENT APPLICATION NUMBER: US/09/370.838  
; CURRENT FILING DATE: 1999-08-09  
; EARLIER APPLICATION NUMBER: US 09/285.323  
; EARLIER FILING DATE: 1999-04-02  
; NUMBER OF SEQ ID NOS: 289  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 119  
; LENGTH: 642  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-370-838-119

Query Match	90.9%	Score 80;	DB 3;	Length 642;
Best Local Similarity	94.3%	Pred. No. 3.6e-20;		
Matches	83;	Conservative	0;	Mismatches 5; Indels 0; Gaps 0;
Qy	1	CGGAGTCGAGTCCTGATCAGCCCTTTTCTACAGCTTTTCTACAAATGGCATTCAAT	60	
Db	535	CGGAGTCGAGTCCTGATCAGCCCTTTTCTACAGCTTTTCTACAAATGGCATTCAAT	594	
Qy	61	AAATGTCAGCTGTTTCTTGGTACAAAAA	88	
Db	595	AAATGTCAGCTGTTTCTTGGTACAAAAA	622	

RESULT 2  
US-09-854-133-119  
; Sequence 119, Application US/09854133  
; Patent No. 6759508  
; GENERAL INFORMATION:  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Mohamath, Roadoh  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Benson, Darin R.  
; APPLICANT: Secrist, Heather

;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR  
;; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER  
;; FILE REFERENCE: 210121.475C10  
;; CURRENT APPLICATION NUMBER: US/09/854.133  
;; CURRENT FILING DATE: 2001-05-11  
;; NUMBER OF SEQ ID NOS: 735  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO 119  
;; LENGTH: 642  
;; TYPE: DNA  
;; ORGANISM: Homo sapien  
US-09-854-133-119

Query Match 90.9%; Score 80; DB 3; Length 642;  
Best Local Similarity 94.3%; Pred. No. 3.6e-20;  
Matches 83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCCTCAGCGCTTTTCTACAATGGCATTCAAT 60  
DB 535 CGGAGCGAGTCCTGTATCAGCCCTTTATCCTCAGCGCTTTTCTACAATGGCATTCAAT 594

QY 61 AAAGTCACGCTGTTCTTGTGTACAAAA 88  
DB 595 AAAGTCACGCTGTTCTTGTGTAAAAA 622

RESULT 3  
US-09-023-655-1279  
; Sequence 1279, Application US/09023655  
; Patent No. 6607879  
; GENERAL INFORMATION:  
; APPLICANT: Cocks, Benjamin G.  
; APPLICANT: Susan G. Stuart  
; APPLICANT: Jeffrey J. Seilhamer  
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE  
; TITLE OF INVENTION: EXPRESSION  
; NUMBER OF SEQUENCES: 1508  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 PORTER DRIVE  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/023.655  
; FILING DATE: HERewith  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Zeller, Karen J.  
; REGISTRATION NUMBER: 37,071  
; REFERENCE/DOCKET NUMBER: PA-0001 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (650) 855-0555  
; TELEFAX: (650) 845-4166  
; INFORMATION FOR SEQ ID NO: 1279:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 808 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: GENBANK  
; CLONE: g311374

US-09-023-655-1279

Query Match 86.1%; Score 75.8; DB 3; Length 808;  
Best Local Similarity 97.5%; Pred. No. 1.5e-18;  
Matches 77; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCCTCAGCGCTTTTCTACAATGGCATTCAAT 60  
DB 722 CGGAGCGAGTCCTGTATCAGCCCTTTATCCTCAGCGCTTTTCTACAATGGCATTCAAT 781

QY 61 AAAGTCACGCTGTTCTTGT 79  
DB 782 AAAGTCACGCTGTTCTTGTGG 800

RESULT 4  
US-09-461-912A-25  
; Sequence 125, Application US/09461912A  
; Patent No. 6709855  
; GENERAL INFORMATION:  
; APPLICANT: Stanton, Lawrence A.  
; APPLICANT: White, R. Tyler  
; APPLICANT: Deam, Deborah L.  
; APPLICANT: Lewicki, John A.  
; TITLE OF INVENTION: Methods for detection and use of  
; TITLE OF INVENTION: differentially expressed genes in disease states  
; FILE REFERENCE: SCIOS.011A  
; CURRENT APPLICATION NUMBER: US/09/461,912A  
; PRIOR FILING DATE: 1999-12-15  
; PRIOR APPLICATION NUMBER: US 60/113,008  
; PRIOR FILING DATE: 1998-12-18  
; NUMBER OF SEQ ID NOS: 48  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 25  
; LENGTH: 808  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (238)...(639)  
US-09-461-912A-25

Query Match 86.1%; Score 75.8; DB 3; Length 808;  
Best Local Similarity 97.5%; Pred. No. 1.5e-18;  
Matches 77; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCCTCAGCGCTTTTCTACAATGGCATTCAAT 60  
DB 722 CGGAGCGAGTCCTGTATCAGCCCTTTATCCTCAGCGCTTTTCTACAATGGCATTCAAT 781

QY 61 AAAGTCACGCTGTTCTTGT 79  
DB 782 AAAGTCACGCTGTTCTTGTGG 800

RESULT 5  
US-10-131-827-2927  
; Sequence 2927, Application US/10131827  
; Patent No. 6905827  
; GENERAL INFORMATION:  
; APPLICANT: Wohlgemuth, Jay  
; APPLICANT: Fry, Kirk  
; APPLICANT: Woodward, Robert  
; APPLICANT: Ly, Ngoc  
; TITLE OF INVENTION: CHRONIC INFLAMMATORY DISEASES  
; FILE REFERENCE: 506612000120  
; CURRENT APPLICATION NUMBER: US/10/131,827  
; CURRENT FILING DATE: 2002-09-06  
; PRIOR APPLICATION NUMBER: US 10/006,290  
; PRIOR FILING DATE: 2001-10-22  
; PRIOR APPLICATION NUMBER: US 60/296,764  
; PRIOR FILING DATE: 2001-06-08



EARLIER APPLICATION NUMBER: US 08/838,762  
EARLIER FILING DATE: 1997-04-09  
EARLIER APPLICATION NUMBER: PCT/US97/00485  
EARLIER FILING DATE: 1997-01-10  
EARLIER APPLICATION NUMBER: US 08/700,014  
EARLIER FILING DATE: 1996-08-20  
EARLIER APPLICATION NUMBER: US 08/585,392  
EARLIER FILING DATE: 1996-01-01  
NUMBER OF SEQ ID NOS: 312  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 283  
LENGTH: 764  
TYPE: DNA  
ORGANISM: Homo sapien  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)...(764)  
OTHER INFORMATION: n = A,T,C or G  
US-09-289-198-283

Query Match 48.6%; Score 42.8; DB 3; Length 764;  
Best Local Similarity 75.7%; Pred. No. 3.5e-06;  
Matches 53; Conservative 0; Mismatches 17; Indels 0; Gaps 0;  
Qy 19 CAGCCCTTTATCCTCACACGCTTTTCTACAATGGCATTCAATAAGTGCACGTTGTTCTT 78  
Db 72 CAGCAGTTTATACCCACACACCTGCTACAGTGTCTATTCATTAATAAGTGCACGTTGCG 13  
Qy 79 GGTACAAAAA 88  
Db 12 AAAAAAAAAA 3

RESULT 9  
US-09-429-755-283/c  
Sequence 283, Application US/09429755A  
Patent No. 6656480  
GENERAL INFORMATION:  
APPLICANT: Frudakis, Tony N.  
APPLICANT: Smith, John M.  
APPLICANT: Reed, Steven G.  
APPLICANT: Misher, Lynda  
APPLICANT: Retter, Marc W.  
APPLICANT: Dillon, Davin C.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
TREATMENT AND DIAGNOSIS OF BREAST CANCER  
FILE REFERENCE: 210121.419C6  
CURRENT APPLICATION NUMBER: US/09/429.755A  
CURRENT FILING DATE: 1999-10-28  
NUMBER OF SEQ ID NOS: 315  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 283  
LENGTH: 764  
TYPE: DNA  
ORGANISM: Homo sapien  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)...(764)  
OTHER INFORMATION: n = A,T,C or G  
US-09-429-755-283

Query Match 48.6%; Score 42.8; DB 3; Length 764;  
Best Local Similarity 75.7%; Pred. No. 3.5e-06;  
Matches 53; Conservative 0; Mismatches 17; Indels 0; Gaps 0;  
Qy 19 CAGCCCTTTATCCTCACACGCTTTTCTACAATGGCATTCAATAAGTGCACGTTGTTCTT 78  
Db 72 CAGCAGTTTATACCCACACACCTGCTACAGTGTCTATTCATTAATAAGTGCACGTTGCG 13  
Qy 79 GGTACAAAAA 88  
Db 12 AAAAAAAAAA 3

RESULT 10  
US-09-699-295-283/c  
Sequence 283, Application US/09699295  
Patent No. 6828431  
GENERAL INFORMATION:  
APPLICANT: Frudakis, Tony N.  
APPLICANT: Reed, Steven G.  
APPLICANT: Smith, John M.  
APPLICANT: Misher, Linda E.  
APPLICANT: Dillon, Davin C.  
APPLICANT: Retter, Marc W.  
APPLICANT: Wang, Aijun  
APPLICANT: Skeiky, Yaseir A.W.  
APPLICANT: Harlocker, Susan L.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
TREATMENT AND DIAGNOSIS OF BREAST CANCER  
FILE REFERENCE: 210121.419C10  
CURRENT APPLICATION NUMBER: US/09/699,295  
CURRENT FILING DATE: 2000-10-26  
NUMBER OF SEQ ID NOS: 326  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 283  
LENGTH: 764  
TYPE: DNA  
ORGANISM: Homo sapien  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)...(764)  
OTHER INFORMATION: n = A,T,C or G  
US-09-699-295-283

Query Match 48.6%; Score 42.8; DB 3; Length 764;  
Best Local Similarity 75.7%; Pred. No. 3.5e-06;  
Matches 53; Conservative 0; Mismatches 17; Indels 0; Gaps 0;  
Qy 19 CAGCCCTTTATCCTCACACGCTTTTCTACAATGGCATTCAATAAGTGCACGTTGTTCTT 78  
Db 72 CAGCAGTTTATACCCACACACCTGCTACAGTGTCTATTCATTAATAAGTGCACGTTGCG 13  
Qy 79 GGTACAAAAA 88  
Db 12 AAAAAAAAAA 3

RESULT 11  
US-09-534-825A-283/c  
Sequence 283, Application US/09534825A  
Patent No. 6861506  
GENERAL INFORMATION:  
APPLICANT: Frudakis, Tony N.  
APPLICANT: Smith, John M.  
APPLICANT: Reed, Steven G.  
APPLICANT: Misher, Lynda  
APPLICANT: Retter, Marc W.  
APPLICANT: Dillon, Davin C.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
TREATMENT AND DIAGNOSIS OF BREAST CANCER  
FILE REFERENCE: 210121.419C7  
CURRENT APPLICATION NUMBER: US/09/534,825A  
CURRENT FILING DATE: 2000-03-23  
NUMBER OF SEQ ID NOS: 317  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 283  
LENGTH: 764  
TYPE: DNA  
ORGANISM: Homo sapien  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)...(764)  
OTHER INFORMATION: n = A,T,C or G  
US-09-534-825A-283

```
Query Match      48.6%; Score 42.8; DB 3; Length 764;
Best Local Similarity 75.7%; Pred. No. 3.5e-06;
Matches 53; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 19 CAGCCCTTATCCTCACACGGCTTTTCTACAAATGCGCATCAATAAAGTGCAGGTGTTCTT 78
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 72 CAGCAGTTTATACCCACACACCTGCTCTACAGTGTCAATCAATAAAGTGCAGGTGCTTGGC 13

Qy 79 GGTCAAAAAA 88
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 12 AAAAAAANA 3

RESULT 12
US-09-949-016-13151
; Sequence 13151, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13151
; LENGTH: 25202
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13151

Query Match      31.1%; Score 27.4; DB 3; Length 25202;
Best Local Similarity 59.7%; Pred. No. 6.9;
Matches 46; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

Qy 6 TCGAGTCCTGTATCAGCCCTTTATCCTCACACGGCTTTTCTACAAATGCGCATCAATAAAGT 65
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 122 TAGAATCCTGAGTCAGCTTTTCCCTCCTCTGATTCCTACTCCCTCCCTTAATCAAAAT 181

Qy 66 GCACGTGTTCTTGTA 82
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 182 TCCATTGTTGGTGCA 198

RESULT 13
US-09-702-705-314/c
; Sequence 314, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 314
```

```
; LENGTH: 513
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-702-705-314

Query Match      30.5%; Score 26.8; DB 3; Length 513;
Best Local Similarity 64.5%; Pred. No. 3;
Matches 40; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy 26 TTATCCTCACACGGCTTTTCTACAATGCGCATTTCAATAAAGTGCAGGTGTTCTTGGTACAA 85
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 198 TTTACCTTTCACGATTTCTTATATGTTATTAATAAAGTTACGCAATTTATGGGAGAA 139

Qy 86 AA 87
    ||
Db 138 AA 137

RESULT 14
US-09-736-457-314/c
; Sequence 314, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 314
; LENGTH: 513
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-736-457-314

Query Match      30.5%; Score 26.8; DB 3; Length 513;
Best Local Similarity 64.5%; Pred. No. 3;
Matches 40; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy 26 TTATCCTCACACGGCTTTTCTACAATGCGCATTTCAATAAAGTGCAGGTGTTCTTGGTACAA 85
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 198 TTTACCTTTCACGATTTCTTATATGTTATTAATAAAGTTACGCAATTTATGGGAGAA 139

Qy 86 AA 87
    ||
Db 138 AA 137

RESULT 15
US-09-614-124B-314/c
; Sequence 314, Application US/09614124B
; Patent No. 6630574
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
```

```
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C9
; CURRENT APPLICATION NUMBER: US/09/614,124B
; CURRENT FILING DATE: 2001-07-11
; NUMBER OF SEQ ID NOS: 1668
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 314
; LENGTH: 513
; TYPE: DNA
; ORGANISM: Homo sapien
; US-09-614-124B-314

Query Match      30.5%; Score 26.8; DB 3; Length 513;
Best Local Similarity 64.5%; Pred. No. 3;
Matches 40; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy      26 TTATCCTCACACGCTTTTCTACAAATGGCATTCAATAAAGTGCACGCTGTTTCTTGGTACAA 85
Db      198 TTTACCTTTCACCTGATTTCTATATGGTATATAAATTAAAGTTCAGGCATTTTATGGGGAGAA 139

Qy      86 AA 87
Db      138 AA 137

Search completed: December 6, 2005, 11:47:20
Job time : 40.4091 secs
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 21:43:03 ; Search time 787.105 Seconds  
(without alignments)  
6427.437 Million cell updates/sec

Title: US-09-980-046B-4  
Perfect score: 89  
Sequence: 1 cgggaatcccccttcagttc.....aaaaaaaaaaaaaagaaaaa 89

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 117666282

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

1: gb\_ba.\*  
2: gb\_in.\*  
3: gb\_env.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pr.\*  
9: gb\_ro.\*  
10: gb\_sts.\*  
11: gb\_by.\*  
12: gb\_un.\*  
13: gb\_vl.\*  
14: gb\_htg.\*  
15: gb\_pl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	79.4	89.2	1464	BD270355	BD270355 50 human
2	79.4	89.2	1757	AK223104	AK223104 Homo sapi
3	70	78.7	428	CQ923987	CQ923987 Sequence
c	4	70	78.7	522	CS113124 Sequence
5	70	78.7	592	AX780236	AX780236 Sequence
6	70	78.7	1761	BC017716	BC017716 Homo sapi
7	70	78.7	1766	AK223068	AK223068 Homo sapi
8	70	78.7	2000	CQ776411	CQ776411 Sequence
9	70	78.7	2000	CQ861464	CQ861464 Sequence
10	70	78.7	2000	CQ981630	CQ981630 Sequence
11	70	78.7	2000	AR569106	AR569106 Sequence
12	70	78.7	2000	AX055560	AX055560 Sequence
13	70	78.7	2000	AX333755	AX333755 Sequence
14	70	78.7	2000	HSU77643	U77643 Homo sapien
15	70	78.7	2180	AR071760	AR071760 Sequence
16	70	78.7	2180	AR097657	AR097657 Sequence
17	70	78.7	2180	AR195192	AR195192 Sequence
c	18	70	78.7	71352 14	AC137090 Homo sapi

19	70	78.7	154252	8	AC132938	AC132938	Homo sapi
20	70	78.7	166458	14	AC084737	AC084737	Homo sapi
21	70	78.7	171569	8	AC132872	AC132872	Homo sapi
22	54	60.7	401	6	AX778685	AX778685	Sequence
23	43.6	49.0	177	6	BD112129	BD112129	EST and e
24	43.6	49.0	177	6	AR416576	AR416576	Sequence
25	43.6	49.0	177	6	AX977270	AX977270	Sequence
26	43.6	49.0	161296	2	AC091483	AC091483	Ttynanos
27	43.2	48.5	2720	9	BC058402	BC058402	Mus muscu
28	42.6	47.9	208462	9	AC129179	AC129179	Mus muscu
c	29	42.6	47.9	232320	9	AC127314	Mus muscu
c	30	42.6	47.9	349980	6	AX344565	Sequence
31	42.2	47.4	1329	8	BC053641	BC053641	Homo sapi
32	42.2	47.4	1386	5	BC090110	BC090110	Xenopus t
33	42.2	47.4	1526	8	BC041705	BC041705	Homo sapi
34	41.6	46.7	1647	8	AF218005	AF218005	Homo sapi
35	41.6	46.7	1961	8	AB060916	AB060916	Macaca fa
36	41.6	46.7	3273	9	BC023693	BC023693	Mus muscu
c	37	41.6	46.7	349980	6	AX344570	Sequence
38	41.4	46.5	1817	9	BC051174	BC051174	Mus muscu
39	41.4	46.5	2057	8	BC017236	BC017236	Homo sapi
c	40	41.4	46.5	110000	14	PFMAL13_24	Continuation (25 o
c	41	41.4	46.5	138622	9	AC124401	Mus muscu
42	41.4	46.5	179155	9	AL772401	AL772401	Mouse DNA
c	43	41.4	46.5	208458	9	AC118009	Mus muscu
c	44	41.4	46.5	219132	9	AC151478	Mus muscu
c	45	41.2	46.3	216215	8	HSG256022	Human DNA

#### ALIGNMENTS

RESULT 1	LOCUS	BD270355	50 human secreted proteins.	1464 bp	DNA	linear	PAT 17-JUL-2003
BD270355	DEFINITION	BD270355	50 human secreted proteins.				
	ACCESSION	BD270355					
	VERSION	BD270355.1	GI:33080123				
	KEYWORDS	JP 2002539775-A/44.					
	SOURCE	Homo sapiens (human)					
	ORGANISM	Homo sapiens					
	REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.					
	AUTHORS	1 (bases 1 to 1464)					
	TITLE	Rosen,C.A., Ruben,S.M. and Komatsoulis,G.					
	JOURNAL	50 human secreted proteins					
	COMMENT	Patent: JP 2002539775-A 44 26-NOV-2002; HUMAN GENOME SCIENCES INC					
		OS Homo sapiens (human)					
		PN JP 2002539775-A/44					
		PD 26-NOV-2002					
		PF 09-MAR-2000 JP 2000606612					
		PR 19-MAR-1999 US 60/125360,11-JUN-1999 US 60/138626 PR					
		03-DEC-1999 US 60/168662					
		PI CRAIG A ROSEN,STEVEN M RUBEN,GEORGE KOMATSOUKIS PC					
		C12N15/09,A61K31/711,A61K38/00,A61K48/00,A61P1/00,A61P1/14, PC					
		A61P1/16,					
		PC					
		A61P3/00,A61P3/06,A61P3/10,A61P7/00,A61P7/02,A61P7/06,A61P9/00, PC					
		A61P9/06,					
		PC A61P9/10,A61P9/10,A61P15/00,A61P17/02,A61P17/06,A61P19/00, PC					
		A61P25/00,					
		PC A61P25/14,A61P25/16,A61P25/24,A61P25/28,A61P27/00,A61P27/06,					
		PC A61P29/00,					
		PC A61P31/00,A61P31/04,A61P31/14,A61P31/16,A61P31/18,A61P31/20,					
		PC A61P31/22,					
		PC A61P33/00,A61P33/02,A61P33/06,A61P35/00,A61P35/02,A61P37/00,					
		PC A61P37/06,					
		PC A61P37/08,A61P43/00,C07K14/47,C07K16/18,C12N1/15,C12N1/19, PC					
		C12N1/21,					
		PC C12N5/10,C12P21/02,C12Q1/02,C12Q1/68,G01N33/68// (C12P21/02, PC					
		C12R1:19),					

```

PC      C12N15/00.A61K37/02.C12N5/00
CC      50 human secreted proteins
FH      Key
FT      source
FT      Location/Qualifiers
FEATURES             source
    1..1464
        /organism="Homo sapiens"
        /mol_type="genomic DNA"
        /db_xref="taxon:9606"
ORIGIN
Query Match      89.2%; Score 79.4; DB 6; Length 1464;
Best Local Similarity 93.3%; Pred. No. 1.2e-07;
Matches 83; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Qy      1 CGGGAATCCCCCTTCAGTCTCTTTGAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
Db      1071 CGGGAATCCCCCTTCAGTCTCTTTGAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1130
Qy      61 AACAAACCAAAAAAAAAAAAAAAAAAAAAA 89
Db      1131 AACAAACCGACTCACAAAAAAAAAAAAAAAA 1159
RESULT 2
LOCUS      AK223104 1757 bp mRNA linear PRI 26-APR-2005
DEFINITION Homo sapiens mRNA for secreted and transmembrane 1 precursor
            variant, clone: KAT04885.
ACCESSION      AK223104
VERSION      AK223104.1 GI:62897768
KEYWORDS      FLI_CDNA; oligo capping.
SOURCE      Homo sapiens
ORGANISM      Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominiidae; Homo.
REFERENCE      1
AUTHORS      Maruyama,K. and Sugano,S.
TITLE      Oligo-capping: a simple method to replace the cap structure of
            eukaryotic mRNAs with oligoribonucleotides
JOURNAL      Gene 138 (1-2), 171-174 (1994)
PUBMED      8125298
REFERENCE      2
AUTHORS      Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and
            Sugano,S.
TITLE      Construction and characterization of a full length-enriched and a
            5'-end-enriched cDNA library
JOURNAL      Gene 200 (1-2), 149-156 (1997)
PUBMED      9373149
REFERENCE      3 (bases 1 to 1757)
AUTHORS      Suzuki,Y., Sugano,S., Totoki,Y., Toyoda,A., Takeda,T., Sakaki,Y.,
            Tanaka,A. and Yokoyama,S.
TITLE      Direct Submission
JOURNAL      Submitted (22-APR-2005) Akiko Tanaka, RIKEN Yokohama Institute,
            Protein Research Group; 1-7-22 Suehiro, Tsurumi, Yokohama,
            Kanagawa, 230-0945, Japan (E-mail:aktanaka@riken.jp,
            URL:http://protein.gsc.riken.jp/, Tel:81-45-503-9452,
            Fax:81-45-503-9450)
COMMENT      This work was supported in part by the National Project on Protein
            Structural and Functional Analysis, Ministry of Education, Culture,
            Sports, Science and Technology of Japan.
            Sumio Sugano, Yutaka Suzuki
            Laboratory of Functional Genomics Department of Medical Genome
            Sciences Graduate School of Frontier Sciences The University of
            Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639 Japan email:
            ssugano@k.u-tokyo.ac.jp
            URL:http://www.k.u-tokyo.ac.jp/index.html.en.
            Location/Qualifiers
            1..1757
                /organism="Homo sapiens"
                /mol_type="mRNA"
FEATURES             source
    1..1757
        /organism="Homo sapiens"
        /mol_type="mRNA"

```

```

/db_xref="taxon:9606"
/clone="KAT04885"
/cell_line="KATO III"
/clone_lib="KAT"
/notes="cloning vector: pME18SFL3
this clone is also named as hss001002732"
<123..869
/note="Start codon is not identified."
/codon_start=1
/evidence=not_experimental
/product="secreted and transmembrane 1 precursor variant"
/protein_id="BAD96824.1"
/db_xref="GI:62897769"
/translacion="MOTCLAPPGHVSQALGTLFLAASLSAONEGWDSPICTEGVVS
VSGENTVSCNISNFAHSHVNIKLRAHQBSAIFNEVAPGYFDFRQWLQVQGVGAQL
VIRGADSHAGLYMHLVGHQRNNRQVTLVSVGAEPQSDPTDGFMPVPVAVTVFILL
VALVFWYRCRSQORREKFFLLPQMKVAALRAQAQQGLSRASAEALWTPDSEPTP
RPLALVFKPSPLGALLELLSQPLFPYAADP"
CDS
Query Match      89.2%; Score 79.4; DB 8; Length 1757;
Best Local Similarity 93.3%; Pred. No. 1.1e-07;
Matches 83; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Qy      1 CGGGAATCCCCCTTCAGTCTCTTTGAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
Db      1650 CGGGAATCCCCCTTCAGTCTCTTTGAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1709
Qy      61 AACAAACCAAAAAAAAAAAAAAAAAAAAAA 89
Db      1710 AACAAACCGACTCACAAAAAAAAAAAAAAAA 1738
RESULT 3
LOCUS      CQ923987 428 bp DNA linear PAT 23-NOV-2004
DEFINITION Sequence 5187 from Patent WO2004097052.
ACCESSION      CQ923987
VERSION      CQ923987.1 GI:56213928
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominiidae; Homo.
REFERENCE      1
AUTHORS      Burczynski,M.E., Twine,N.C., Slonim,D.K., Trepicchio,W.L.,
            Straha,A., Immerman,F. and Dörner,A.J.
TITLE      Methods for prognosis and treatment of solid tumors
JOURNAL      Patent: WO 2004097052-A 5187 11-NOV-2004;
            Wyeth (US); Burczynski, Michael E. (US)
FEATURES             source
    1..428
        /organism="Homo sapiens"
        /mol_type="unassigned DNA"
        /db_xref="taxon:9606"
    misc_feature      234..240
        /note="n includes a, c, g, or t, or contains no
        nucleotide"
    misc_feature      242..249
        /note="n includes a, c, g, or t, or contains no
        nucleotide"
    misc_feature      252..256
        /note="n includes a, c, g, or t, or contains no
        nucleotide"
ORIGIN
Query Match      78.7%; Score 70; DB 6; Length 428;
Best Local Similarity 93.6%; Pred. No. 1.8e-05;
Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy      1 CGGGAATCCCCCTTCAGTCTCTTTGAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60

```

Db 100 CGGGAATCCCTTCAGTTCCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 159

Qy 61 AACAAACCAAAAAA 78  
 |||||  
 Db 160 AACAAACCGACTCACA 177  
 |||||

RESULT 4  
 LOCUS CS113124 522 bp DNA PAT 24-JUN-2005  
 DEFINITION Sequence 142 from Patent WO2005054507.  
 ACCESSION CS113124  
 VERSION CS113124.1 GI:58224696  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Hominidae; Homo.

REFERENCE 1  
 AUTHORS Corfe, B. and Chirakkal, H.  
 TITLE Gene screen  
 JOURNAL Patent: WO 2005054507-A 142 16-JUN-2005;  
 University of Sheffield (GB)  
 FEATURES source  
 1..522  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"

ORIGIN  
 Query Match 78.7%; Score 70; DB 6; Length 522;  
 Best Local Similarity 93.6%; Pred. No. 1.7e-05;  
 Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCTTCAGTTCCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60  
 |||||  
 Db 82 CGGGAATCCCTTCAGTTCCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 23  
 |||||

Qy 61 AACAAACCAAAAAA 78  
 |||||  
 Db 22 AACAAACCGACTCACA 5  
 |||||

RESULT 5  
 LOCUS AX780236 592 bp DNA PAT 14-JUL-2003  
 DEFINITION Sequence 2393 from Patent WO03039443.  
 ACCESSION AX780236  
 VERSION AX780236.1 GI:32697230  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Hominidae; Homo.

REFERENCE 1  
 AUTHORS Haeflrich, T., Schoch, C., Kern, W., Kohlmann, A., Schnittger, S.,  
 Dugas, M., Eils, R., Brors, B. and Mergenthaler, S.  
 TITLE Novel genetic markers for leukemias  
 JOURNAL Patent: WO 03039443-A 2393 15-MAY-2003;  
 Deutsches Krebsforschungszentrum (DE) ;  
 Ludwig-Maximilian-Universitaet Muenchen (DE) ;  
 PD Dr. Dr. (DE) ; Schoch, Claudia (DE) ; Kern, Wolfgang (DE)  
 FEATURES source  
 1..592  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"

ORIGIN  
 Query Match 78.7%; Score 70; DB 6; Length 592;  
 Best Local Similarity 93.6%; Pred. No. 1.7e-05;

Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCTTCAGTTCCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60  
 |||||  
 Db 434 CGGGAATCCCTTCAGTTCCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 493  
 |||||

Qy 61 AACAAACCAAAAAA 78  
 |||||  
 Db 494 AACAAACCGACTCACA 511  
 |||||

RESULT 6  
 LOCUS BC017716 1761 bp mRNA linear PRI 29-JUN-2004  
 DEFINITION Homo sapiens secreted and transmembrane 1, mRNA (cDNA clone  
 MGC:21265 IMAGE:4413387), complete cds.  
 ACCESSION BC017716  
 VERSION BC017716.1 GI:17389345  
 KEYWORDS MGC.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Hominidae; Homo.

REFERENCE 1 (bases 1 to 1761)  
 AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,  
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,  
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,  
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,  
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.W., Hong, L.,  
 Scapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,  
 Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,  
 Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,  
 Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,  
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,  
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,  
 Villalón, D.C., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,  
 Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,  
 Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,  
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,  
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,  
 Butterfield, Y.S., Krzywinska, M.I., Skalska, U., Smallos, D.E.,  
 Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.  
 Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences  
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
 12477932

2 (bases 1 to 1761)  
 AUTHORS Strausberg, R.  
 TITLE Direct Submission  
 JOURNAL Submitted (03-DEC-2001) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA  
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 Contact: MGC help desk  
 Email: [cgaps-r@mail.nih.gov](mailto:cgaps-r@mail.nih.gov)  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing Arrayed by: The I.M.A.G.E. Consortium (LINL)  
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome  
 Center, Stanford University School of Medicine, Stanford, CA 94305  
 Web site: <http://www-sbgc.stanford.edu>  
 Contact: (Dickson, Mark) [mcd@paxil.stanford.edu](mailto:mcd@paxil.stanford.edu)  
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, J., Rodriguez, A., and Myers,  
 R. M.

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LINL at: <http://image.llnl.gov>  
 Series: IPAK Plate: 27 Row: h Column: 10  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: matched mRNA gi: 4506868.  
 Location/Qualifiers

FEATURES

source	1. .1761 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="MGC:21265 IMAGE:4413387" /tissue_type="Liver, adenocarcinoma" /clone_lib="NIH_MGC_90" /lab_host="DH10B" /note="Vector: pCMV-SPORT6" 1. .1761 /gene="SECTM1" /note="synonym: K12" /db_xref="GeneID:6398" /db_xref="MIM:602602" 93. .839 /gene="SECTM1" /codon_start=1 /product="secreted and transmembrane 1 precursor" /protein_id="AAH17716.1" /db_xref="GI:17389346" /db_xref="GeneID:6398" /db_xref="MIM:602602" /translation="MOTCPLAFFGHVSQALGTLFLAASLSAQNEGWDSPICTEGVVS VSGENTVMSCNISNFSHVNIKLRAHQESAIENEVAPGYFRDGHQLOVQGVQVL VIKGARDSHAGLWMHLVGHORNNROVTLVSGAEPOSADPTGFWPVPVAVTVFILL VALVMPAWYRCRCQQRRKKFFLLBPQMKFAALRAGAQQGLSRASAEIWPDPSEPTP RPLALVFKPSPLGALLELLSQPLFPFYAADP"
gene	
CDS	
FEATURES	source 1. .1766 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="KAT02983" /cell_line="KATO III" /clone_lib="KAT" /notes="cloning vector: pME18SFL3 this clone is also named as hss001002587" <92. .838 /notes="Start codon is not identified." /codon_start=1 /evidence="not_experimental /product="secreted and transmembrane 1 precursor variant" /protein_id="BAD96788.1" /db_xref="GI:62897697" /translation="MOTCPLAFFGHVSQALGTLFLAASLSAQNEGWDSPICTEGVVS VSGENTVMSCNISNFSHVNIKLRAHQESAIENEVAPGYFRDGHQLOVQGVQVL VIKGARDSHAGLWMHLVGHORNNROVTLVSGAEPOSADPTGFWPVPVAVTVFILL VALVMPAWYRCRCQQRRKKFFLLBPQMKFAALRAGAQQGLSRASAEIWPDPSEPTP RPLALVFKPSPLGALLELLSQPLFPFYAADP"
CDS	
ORIGIN	Query Match 78.7%; Score 70; DB 8; Length 1761; Best Local Similarity 93.6%; Pred. No. 1.2e-05; Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  Qy 1 CGGGAATCCCCCTTCAGTCTTTGAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60       Db 1619 CGGGAATCCCCCTTCAGTCTTTGAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1678  Qy 61 AACAAACCAAAAAA 78       Db 1679 AACAAACCGACTCACAAA 1696  RESULT 8 CQ776411 2000 bp DNA linear PAT 11-MAR-2004 LOCUS DEFINITION Sequence 97 from Patent EP1394274. ACCESSION CQ776411 VERSION CQ776411.1 GI:45379801 KEYWORDS SOURCE Homo sapiens ORGANISM Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo. 1 REFERENCE AUTHORS TITLE Methods of testing for bronchial asthma or chronic obstructive pulmonary disease JOURNAL Patent: EP 1394274-A 97 03-MAR-2004; Genox Research, Inc. (JP) FEATURES source 1. .2000 /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606"
source	1. .1761 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="MGC:21265 IMAGE:4413387" /tissue_type="Liver, adenocarcinoma" /clone_lib="NIH_MGC_90" /lab_host="DH10B" /note="Vector: pCMV-SPORT6" 1. .1761 /gene="SECTM1" /note="synonym: K12" /db_xref="GeneID:6398" /db_xref="MIM:602602" 93. .839 /gene="SECTM1" /codon_start=1 /product="secreted and transmembrane 1 precursor" /protein_id="AAH17716.1" /db_xref="GI:17389346" /db_xref="GeneID:6398" /db_xref="MIM:602602" /translation="MOTCPLAFFGHVSQALGTLFLAASLSAQNEGWDSPICTEGVVS VSGENTVMSCNISNFSHVNIKLRAHQESAIENEVAPGYFRDGHQLOVQGVQVL VIKGARDSHAGLWMHLVGHORNNROVTLVSGAEPOSADPTGFWPVPVAVTVFILL VALVMPAWYRCRCQQRRKKFFLLBPQMKFAALRAGAQQGLSRASAEIWPDPSEPTP RPLALVFKPSPLGALLELLSQPLFPFYAADP"
gene	
CDS	
FEATURES	source 1. .1766 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="KAT02983" /cell_line="KATO III" /clone_lib="KAT" /notes="cloning vector: pME18SFL3 this clone is also named as hss001002587" <92. .838 /notes="Start codon is not identified." /codon_start=1 /evidence="not_experimental /product="secreted and transmembrane 1 precursor variant" /protein_id="BAD96788.1" /db_xref="GI:62897697" /translation="MOTCPLAFFGHVSQALGTLFLAASLSAQNEGWDSPICTEGVVS VSGENTVMSCNISNFSHVNIKLRAHQESAIENEVAPGYFRDGHQLOVQGVQVL VIKGARDSHAGLWMHLVGHORNNROVTLVSGAEPOSADPTGFWPVPVAVTVFILL VALVMPAWYRCRCQQRRKKFFLLBPQMKFAALRAGAQQGLSRASAEIWPDPSEPTP RPLALVFKPSPLGALLELLSQPLFPFYAADP"
CDS	
ORIGIN	Query Match 78.7%; Score 70; DB 8; Length 1761; Best Local Similarity 93.6%; Pred. No. 1.2e-05; Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  Qy 1 CGGGAATCCCCCTTCAGTCTTTGAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60       Db 1620 CGGGAATCCCCCTTCAGTCTTTGAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1679  Qy 61 AACAAACCAAAAAA 78       Db 1680 AACAAACCGACTCACAAA 1697  RESULT 7 AK223068 1766 bp mRNA linear PRI 26-APR-2005 LOCUS DEFINITION Homo sapiens mRNA for secreted and transmembrane 1 precursor variant, clone: KAT02983. ACCESSION AK223068 VERSION AK223068.1 GI:62897696 KEYWORDS FLI_CDNA; oligo capping. SOURCE Homo sapiens ORGANISM Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo. 1 REFERENCE AUTHORS Maruyama,K. and Sugano,S. TITLE Oligo-capping: A simple method to replace the cap structure of eukaryotic mRNAs with oligoribonucleotides JOURNAL Gene 138 (1-2), 171-174 (1994) PUBMED 8125298 REFERENCE AUTHORS Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano,S. TITLE Construction and characterization of a full length-enriched and a 5'-end-enriched cDNA library JOURNAL Gene 200 (1-2), 149-156 (1997) PUBMED 9373149 REFERENCE AUTHORS Suzuki,Y., Sugano,S., Totoki,Y., Toyoda,A., Takeda,T., Sakaki,Y., Tanaka,A. and Yokoyama,S. TITLE Direct Submission JOURNAL Submitted (22-APR-2005) Akiko Tanaka, RIKEN Yokohama Institute,

```
ORIGIN
Query Match 78.7%; Score 70; DB 6; Length 2000;
Best Local Similarity 93.6%; Pred. No. 1.1e-05;
Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGAATCCCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
Db 1645 CGGAATCCCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1704

Qy 61 AACAAACCAAAAAA 78
Db 1705 AACAAACCGACTCACAA 1722

RESULT 9
LOCUS CQ861464 2000 bp DNA linear PAT 10-SEP-2004
DEFINITION Sequence 97 from Patent WO2004072265.
ACCESSION CQ861464
VERSION CQ861464.1 GI:51982453
KEYWORDS Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1
REFERENCE
AUTHORS Burczynski, M., Twine, N., Dorner, A. J. and Trepicchio, W. L.
TITLE METHODS FOR MONITORING DRUG ACTIVITIES IN VIVO / I
JOURNAL Patent: WO 2004072265-A 97 26-AUG-2004;
Weth (US); Burczynski, Michael E. (US); Twine, Natalie C. (US);
Dorner, Andrew J. (US); Trepicchio, William L. (US)
FEATURES
source Location/Qualifiers
1. .2000
/mol_type="Homo sapiens"
/db_xref="taxon:9606"

ORIGIN
Query Match 78.7%; Score 70; DB 6; Length 2000;
Best Local Similarity 93.6%; Pred. No. 1.1e-05;
Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGAATCCCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
Db 1645 CGGAATCCCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1704

Qy 61 AACAAACCAAAAAA 78
Db 1705 AACAAACCGACTCACAA 1722

RESULT 10
LOCUS CQ981630 2000 bp DNA linear PAT 25-JAN-2005
DEFINITION Sequence 485 from Patent EP1498424.
ACCESSION CQ981630
VERSION CQ981630.1 GI:58190920
KEYWORDS Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1
REFERENCE
AUTHORS Rosenthal, A., Hermann, K., Heiden, K., Pilarsky, C., Bruesmendorf, T.,
Staub, E., Roepcke, S., Mennerich, D., Kinnemann, H. and Li, X.
TITLE Human nucleic acid sequences from lung tumours
JOURNAL Patent: EP 1498424-A 485 19-JAN-2005;
Hinzmann, Bernd (DE); Hermann, Klaus (DE); Heiden, Esmeralda (DE);
Rosenthal, Andre (DE)
FEATURES
source Location/Qualifiers
1. .2000
/mol_type="Homo sapiens"
/db_xref="taxon:9606"

ORIGIN
Query Match 78.7%; Score 70; DB 6; Length 2000;
Best Local Similarity 93.6%; Pred. No. 1.1e-05;
Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGAATCCCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
Db 1645 CGGAATCCCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1704

Qy 61 AACAAACCAAAAAA 78
Db 1705 AACAAACCGACTCACAA 1722

RESULT 11
LOCUS AR569106 2000 bp DNA linear PAT 14-DEC-2004
DEFINITION Sequence 3 from patent US 6762030.
ACCESSION AR569106
VERSION AR569106.1 GI:56569141
KEYWORDS Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 2000)
Lyman, S. D. and Fanslow, W. C. III.
Ligand for CD7 and methods for use thereof
Patent: US 6762030-A 3 13-JUL-2004;
Immunex Corporation; Seattle, WA;
WOX;
FEATURES
source Location/Qualifiers
1. .2000
/mol_type="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 78.7%; Score 70; DB 6; Length 2000;
Best Local Similarity 93.6%; Pred. No. 1.1e-05;
Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGAATCCCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
Db 1645 CGGAATCCCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1704

Qy 61 AACAAACCAAAAAA 78
Db 1705 AACAAACCGACTCACAA 1722

RESULT 12
LOCUS AX055560 2000 bp DNA linear PAT 13-JAN-2001
DEFINITION Sequence 3 from Patent WO0073333.
ACCESSION AX055560
VERSION AX055560.1 GI:12228778
KEYWORDS Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1
REFERENCE
AUTHORS Lyman, S. D. and Fanslow, W. C.
TITLE Ligand for cd7, and methods for use thereof
JOURNAL Patent: WO 0073333-A 3 07-DEC-2000;
Immunex Corporation (US)
FEATURES
source Location/Qualifiers
1. .2000
/mol_type="Homo sapiens"
/db_xref="taxon:9606"

ORIGIN
Query Match 78.7%; Score 70; DB 6; Length 2000;
Best Local Similarity 93.6%; Pred. No. 1.1e-05;
Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGAATCCCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
Db 1645 CGGAATCCCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1704

Qy 61 AACAAACCAAAAAA 78
Db 1705 AACAAACCGACTCACAA 1722
```

```
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
119..865
/codon_start=1
/notes="unnamed protein product"
/db_xref="GI:12228779"
/translation="MQTCLAPFGHVSQALGTLFLAASLSAQNCGWDSICTEGVVS
VSGENTVMSCNISAFSHVNIKLRAHQESALFNEVAPGYFRDQGLQVQGGVAQL
VIKGRDASHAGLYMHVGHQRNRRQVTLVSGAEPQSDPTGFWPVPVAVTAVFILL
VALVMPAWYRCRCSQQRREKFFLEPQMKVAALRAGAQQGLSRASAEIWTDPSEPTP
RPLALVPKPSPLGALLELSPQLFPFYAADP"

ORIGIN
Query Match 78.7%; Score 70; DB 6; Length 2000;
Best Local Similarity 93.6%; Pred. No. 1.1e-05;
Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCTTCAGTCTTTTGAAAAGTTCATGACTCGAATATCTGAATGAAGAA 60
Db 1645 CGGGAATCCCTTCAGTCTTTTGAAAAGTTCATGACTCGAATATCTGAATGAAGAA 1704

Qy 61 AACAAACCAAAAAA 78
Db 1705 AACAAACCGACTCACA 1722

RESULT 13
AX333755 2000 bp DNA linear PAT 09-JAN-2002
LOCUS
DEFINITION Sequence 4264 from Patent WO0194629.
ACCESSION AX333755
VERSION AX333755.1 GI:18124474
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
Horrikan, S., Soppet, D.R. and Weaver, Z.
TITLE Cancer gene determination and therapeutic screening using signature
gene sets
JOURNAL Patent: WO 0194629-A 4264 13-DEC-2001;
Avalon Pharmaceuticals (US)
FEATURES
source
Location/Qualifiers
1..2000
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 78.7%; Score 70; DB 6; Length 2000;
Best Local Similarity 93.6%; Pred. No. 1.1e-05;
Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCTTCAGTCTTTTGAAAAGTTCATGACTCGAATATCTGAATGAAGAA 60
Db 1645 CGGGAATCCCTTCAGTCTTTTGAAAAGTTCATGACTCGAATATCTGAATGAAGAA 1704

Qy 61 AACAAACCAAAAAA 78
Db 1705 AACAAACCGACTCACA 1722

RESULT 14
HSU77643 2000 bp mRNA linear PRI 10-MAR-1998
LOCUS
DEFINITION Homo sapiens K12 protein precursor mRNA, complete cds.
AUTHORS HSU77643
TITLE U77643
VERSION U77643.1 GI:2062390

/organism="Homo sapiens (human)"
/mol_type="mRNA"
/db_xref="taxon:9606"
119..865
/chromosome="17"
/map="17q25.2-25.3"
/cell_type="K562 erythroleukemic"
119..865
/notes="type Ia transmembrane protein"
/codon_start=1
/product="K12 protein precursor"
/protein_id="AAC52044.1"
/db_xref="GI:2062391"
/translation="MQTCLAPFGHVSQALGTLFLAASLSAQNCGWDSICTEGVVS
VSGENTVMSCNISAFSHVNIKLRAHQESALFNEVAPGYFRDQGLQVQGGVAQL
VIKGRDASHAGLYMHVGHQRNRRQVTLVSGAEPQSDPTGFWPVPVAVTAVFILL
VALVMPAWYRCRCSQQRREKFFLEPQMKVAALRAGAQQGLSRASAEIWTDPSEPTP
RPLALVPKPSPLGALLELSPQLFPFYAADP"
119..202
203..862
/product="K12 protein"
551..623
/notes="encodes transmembrane domain"
752..862
/notes="encodes proline-rich region"

ORIGIN
Query Match 78.7%; Score 70; DB 8; Length 2000;
Best Local Similarity 93.6%; Pred. No. 1.1e-05;
Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCTTCAGTCTTTTGAAAAGTTCATGACTCGAATATCTGAATGAAGAA 60
Db 1645 CGGGAATCCCTTCAGTCTTTTGAAAAGTTCATGACTCGAATATCTGAATGAAGAA 1704

Qy 61 AACAAACCAAAAAA 78
Db 1705 AACAAACCGACTCACA 1722

RESULT 15
AR071760 2180 bp DNA linear PAT 18-FEB-2000
LOCUS
DEFINITION Sequence 2 from patent US 5912142.
ACCESSION AR071760
VERSION AR071760.1 GI:7222648
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2180)
AUTHORS Kaufman, R.E. and Slentz-Kesler, K.
TITLE Gene product over expressed in cancer cells
JOURNAL Patent: US 5912142-A 2 15-JUN-1999;
```

FEATURES  
source Location/Qualifiers  
1..2180  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN

Query Match 78.7%; Score 70; DB 6; Length 2180;  
Best Local Similarity 93.6%; Pred. No. 1.1e-05;  
Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
Qy 1 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60  
Db 1841 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1900  
Qy 61 AACAAACCAAAAAAAA 78  
Db 1901 AACAAACCGACTCACAA 1918

Search completed: December 7, 2005, 04:21:58  
Job time : 789.105 secs

*This Page Blank (uspto)*



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 21:43:03 ; Search time 698.666 Seconds  
(without alignments)  
6427.437 Million cell updates/sec

Title: US-09-980-046B-3  
Perfect score: 79  
Sequence: 1 cgggaatcccccttcagttc.....aaacaaaccgactctaaaaa 79

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

#### Database :

GenEmbl.\*

1: gb\_ba.\*  
2: gb\_in.\*  
3: gb\_env.\*  
4: gb\_on.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pr.\*  
9: gb\_ro.\*  
10: gb\_sts.\*  
11: gb\_sy.\*  
12: gb\_un.\*  
13: gb\_vl.\*  
14: gb\_htg.\*  
15: gb\_pl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	75.8	95.9	1464	6	BD270355 50 human
2	75.8	95.9	1757	8	AK223104 Homo sapi
3	74.8	94.7	428	6	CQ223987 Sequence
c	74.8	94.7	522	6	CS113124 Sequence
	74.8	94.7	592	6	AX780236 Sequence
6	74.8	94.7	1761	8	BC017716 Homo sapi
7	74.8	94.7	1766	8	AK223068 Homo sapi
8	74.8	94.7	2000	6	CQ776411 Sequence
9	74.8	94.7	2000	6	CQ861464 Sequence
10	74.8	94.7	2000	6	CQ981630 Sequence
11	74.8	94.7	2000	6	AR569106 Sequence
12	74.8	94.7	2000	6	AX055560 Sequence
13	74.8	94.7	2000	6	AX333755 Sequence
14	74.8	94.7	2000	8	HSU77643 Homo sapien
15	74.8	94.7	2180	6	AR071760 Sequence
16	74.8	94.7	2180	6	AR097657 Sequence
17	74.8	94.7	2180	6	AR195192 Sequence
c	74.8	94.7	71352	14	AC137090 Homo sapi

74.8	94.7	154252	8	AC132938	AC132938	Homo sapi
74.8	94.7	166458	14	AC084737	AC084737	Homo sapi
74.8	94.7	171569	8	AC132872	AC132872	Homo sapi
54	68.4	401	6	AX778685	AX778685	Sequence
34.2	43.3	174294	14	AC013659	AC013659	Homo sapi
34.2	43.3	186783	8	AC090355	AC090355	Homo sapi
34.2	43.3	187167	14	AP001563	AP001563	Homo sapi
33.8	42.8	161296	2	AC091483	AC091483	Trypanosoma
33.4	42.3	178524	8	AC055120	AC055120	Homo sapi
33.4	42.3	191840	8	AC092656	AC092656	Homo sapi
32.2	40.8	55672	14	AC100892	AC100892	Mus muscu
32.2	40.8	152306	9	AC132290	AC132290	Mus muscu
32.2	40.8	226007	14	AC153360	AC153360	Mus muscu
32.2	40.8	255667	9	AC161426	AC161426	Mus muscu
31.8	40.3	205793	14	AC126989	AC126989	Rattus no
31.8	40.3	224158	14	AC093984	AC093984	Rattus no
31.8	40.3	234778	9	AC115778	AC115778	Mus muscu
31.8	40.3	277424	14	AC120729	AC120729	Rattus no
31.2	39.5	167585	8	AL137017	AL137017	Human DNA
30.8	39.0	208462	9	AC129179	AC129179	Mus muscu
30.8	39.0	232320	9	AC127314	AC127314	Mus muscu
30.6	38.7	162281	8	AC105395	AC105395	Homo sapi
30.6	38.7	189631	14	AC009920	AC009920	Homo sapi
30.6	38.7	191935	14	AC114747	AC114747	Homo sapi
30.6	38.7	196318	14	AP001950	AP001950	Homo sapi
30.6	38.7	203114	14	AC011818	AC011818	Homo sapi
30.6	38.7	206364	14	CT009494	CT009494	Mus muscu

#### ALIGNMENTS

RESULT 1	BD270355	50 human secreted proteins.	1464 bp	DNA	linear	PAT 17-JUL-2003
LOCUS	BD270355	50 human secreted proteins.				
DEFINITION	BD270355					
ACCESSION	BD270355					
VERSION	BD270355.1	GI:33080123				
KEYWORDS	JP 2002539775-A/44.					
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.					
AUTHORS	1 (bases 1 to 1464)					
TITLE	Rosen,C.A., Ruben,S.M. and Komatsoulis,G.					
JOURNAL	50 human secreted proteins					
COMMENT	Patent: JP 2002539775-A 44 26-NOV-2002; HUMAN GENOME SCIENCES INC OS Homo sapiens (human) PN JP 2002539775-A/44 PD 26-NOV-2002 PF 09-WAR-2000 JP 2000606612 PR 19-MAR-1999 US 60/125360,11-JUN-1999 US 60/138626 PR 03-DEC-1999 US 60/168662 PI CRAIG A ROSEN,STEVEN M RUBEN,GEORGE KOMATSOUKIS PC C12N15/09,A61K31/711,A61K38/00,A61K48/00,A61P1/00,A61P1/14, PC A61P1/16, PC A61P3/00,A61P3/06,A61P3/10,A61P7/00,A61P7/02,A61P7/06,A61P9/00, PC A61P9/06, PC A61P9/10,A61P9/10,A61P15/00,A61P17/02,A61P17/06,A61P19/00, PC A61P25/00, PC A61P25/14,A61P25/16,A61P25/24,A61P25/28,A61P27/00,A61P27/06, PC A61P29/00, PC A61P31/00,A61P31/04,A61P31/14,A61P31/16,A61P31/18,A61P31/20, PC A61P31/22, PC A61P33/00,A61P33/02,A61P33/06,A61P35/00,A61P35/02,A61P37/00, PC A61P37/06, PC A61P37/08,A61P43/00,C07K14/47,C07K16/18,C12N1/15,C12N1/19, PC C12N1/21, PC C12N5/10,C12P21/02,C12Q1/02,C12Q1/68,G01N33/68//C12P21/02, PC C12R1/19),					

PC C12N15/00.A61K37/02.C12N5/00  
CC 50 human secreted proteins  
FH Key Location/Qualifiers  
FT source 1..1464  
/organism="Homo sapiens (human)".  
FEATURES  
source Location/Qualifiers  
1..1464  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 95.9%; Score 75.8; DB 6; Length 1464;  
Best Local Similarity 97.5%; Pred. No. 1.5e-11;  
Matches 77; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 CGGGAATCCCCCTTCAGTCTTTGAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60  
Db 1071 CGGGAATCCCCCTTCAGTCTTTGAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1130  
Qy 61 AACAAACCGACTCTAAAAA 79  
Db 1131 AACAAACCGACTCACAATA 1149  
RESULT 2  
AK223104 1757 bp mRNA linear PRI 26-APR-2005  
LOCUS Homo sapiens mRNA for secreted and transmembrane 1 precursor  
DEFINITION variant, clone: KAT04885.  
ACCESSION AK223104  
VERSION AK223104.1 GI:62897768  
KEYWORDS FLI\_CDNA; oligo capping.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE  
1 Maruyama,K. and Sugano,S.  
Oligo-capping: a simple method to replace the cap structure of  
eukaryotic mRNAs with oligoribonucleotides  
Gene 138 (1-2), 171-174 (1994)  
JOURNAL PUBMED 8125298  
REFERENCE  
2 Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and  
Sugano,S.  
Construction and characterization of a full length-enriched and a  
5'-end-enriched cDNA library  
Gene 200 (1-2), 149-156 (1997)  
JOURNAL PUBMED 9373149  
REFERENCE  
3 (bases 1 to 1757)  
Suzuki,Y., Sugano,S., Totoki,Y., Toyoda,A., Takeda,T., Sakaki,Y.,  
Tanaka,A. and Yokoyama,S.  
Direct Submission  
Submitted (22-APR-2005) Akiko Tanaka, RIKEN Yokohama Institute,  
Protein Research Group; 1-7-22 Suehiro, Tsurumi, Yokohama,  
Kanagawa, 230-0045, Japan (E-mail:aktanaka@riken.jp,  
URL:http://protein.gsc.riken.jp/, Tel:81-45-503-9452,  
Fax:81-45-503-9450)  
This work was supported in part by the National Project on Protein  
Structural and Functional Analysis, Ministry of Education, Culture,  
Sports, Science and Technology of Japan.  
Sumio Sugano, Yutaka Suzuki  
Laboratory of Functional Genomics Department of Medical Genome  
Sciences Graduate School of Frontier Sciences The University of  
Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639 Japan email:  
ssugano@k.u-tokyo.ac.jp  
URL:http://www.k.u-tokyo.ac.jp/index.html.en.  
Location/Qualifiers  
1..1757  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
FEATURES  
source

/db\_xref="taxon:9606"  
/clone="KAT04885"  
/cell\_line="KATO III"  
/clone\_lib="KAT"  
/notes="cloning vector: pME18SFL3  
this clone is also named as hss001002732"  
<123..869  
/notes="Start codon is not identified."  
/codon\_start=1  
/evidence="not experimental  
/product="secreted and transmembrane 1 precursor variant"  
/protein\_id="BAD96824.1"  
/db\_xref="GI:62897769"  
/translation="MQTCPLAPPGHVSOALGTLPLAASLSAQNEGWDSPICTEGVVS  
VNGENTVNSCNISAFSHVNIKLRAHGQESALFNEVAPGYFSDGWLQVQGGVAQL  
VIKARDSHAGLIMHVLHQNNRQVTLVSGAEQPSAPDTGFWPFAVVTAFVILL  
VALVFWAYRCRRSQORREKFFLLEPQMKVAALRAGAQQGLSRASAEIWTDPSEPTF  
RPLALVFKPSPLGALELLSPQLFPFYAADP"  
CDS  
Query Match 95.9%; Score 75.8; DB 8; Length 1757;  
Best Local Similarity 97.5%; Pred. No. 1.5e-11;  
Matches 77; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 CGGGAATCCCCCTTCAGTCTTTGAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60  
Db 1650 CGGGAATCCCCCTTCAGTCTTTGAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1709  
Qy 61 AACAAACCGACTCTAAAAA 79  
Db 1710 AACAAACCGACTCACAATA 1728  
RESULT 3  
CQ923987 428 bp DNA linear PAT 23-NOV-2004  
LOCUS Sequence 5187 from Patent WO2004097052.  
DEFINITION CQ923987  
ACCESSION CQ923987  
VERSION CQ923987.1 GI:56213928  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE  
1 Burczynski,M.E., Twine,N.C., Slonim,D.K., Trepicchio,W.L.,  
Strahs,A., Immerman,F. and Dörner,A.J.  
Methods for prognosis and treatment of solid tumors  
Patent: WO 2004097052-A 5187 11-NOV-2004;  
Wyeth (US); Burczynski, Michael E. (US)  
FEATURES  
source Location/Qualifiers  
1..428  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
misc\_feature 234..240  
/notes="n includes a, c, g, or t, or contains no  
nucleotide"  
misc\_feature 242..249  
/notes="n includes a, c, g, or t, or contains no  
nucleotide"  
misc\_feature 252..256  
/notes="n includes a, c, g, or t, or contains no  
nucleotide"  
ORIGIN  
Query Match 94.7%; Score 74.8; DB 6; Length 428;  
Best Local Similarity 97.4%; Pred. No. 3.9e-11;  
Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 CGGGAATCCCCCTTCAGTCTTTGAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60  
|||||

Db 100 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 159

QY 61 AACAAACCGACTCTAAAA 78  
|||||

Db 160 AACAAACCGACTCACA 177  
|||||

RESULT 4  
LOCUS CS113124 522 bp DNA linear PAT 24-JUN-2005  
DEFINITION Sequence 142 from Patent WO2005054507.  
ACCESSION CS113124  
VERSION CS113124.1 GI:68224696  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.

REFERENCE 1  
AUTHORS Corfe B. and Chirakkal H.  
TITLE Gene screen  
JOURNAL Patent: WO 2005054507-A 142 16-JUN-2005;  
UNIVERSITY University of Sheffield (GB)  
LOCATION/Qualifiers  
FEATURES  
source 1..522  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 94.7%; Score 74.8; DB 6; Length 522;  
Best Local Similarity 97.4%; Pred. No. 3.8e-11;  
Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60  
|||||

Db 82 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 23  
|||||

QY 61 AACAAACCGACTCTAAAA 78  
|||||

Db 22 AACAAACCGACTCACA 5  
|||||

RESULT 5  
LOCUS AX780236 592 bp DNA linear PAT 14-JUL-2003  
DEFINITION Sequence 2393 from Patent WO03039443.  
ACCESSION AX780236  
VERSION AX780236.1 GI:32697230  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.

REFERENCE 1  
AUTHORS Haerlach, T., Schoch, C., Kern, W., Kohlmann, A., Schnittger, S.,  
Dugas, M., Ellis, R., Brors, B. and Mergenthaler, S.  
TITLE Novel genetic markers for leukemias  
JOURNAL Patent: WO 03039443-A 2393 15-MAY-2003;  
Deutsches Krebsforschungszentrum (DE) ;  
Ludwig-Maximilians-Universitaet Muenchen (DE) ;  
PD Dr. Dr. (DE) ; Schoch, Claudia (DE) ; Kern, Wolfgang (DE)  
LOCATION/Qualifiers  
FEATURES  
source 1..592  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 94.7%; Score 74.8; DB 6; Length 592;  
Best Local Similarity 97.4%; Pred. No. 3.7e-11;

Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60  
|||||

Db 434 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 493  
|||||

QY 61 AACAAACCGACTCTAAAA 78  
|||||

Db 494 AACAAACCGACTCACA 511  
|||||

RESULT 6  
LOCUS BC017716 1761 bp mRNA linear PRI 29-JUN-2004  
DEFINITION Homo sapiens secreted and transmembrane 1, mRNA (cDNA clone  
MGC:21265 IMAGE:4413387), complete cds.  
ACCESSION BC017716  
VERSION BC017716 GI:17389345  
KEYWORDS MGC.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.

REFERENCE 1 (bases 1 to 1761)  
AUTHORS Klausner, R.D., Collins, F.S., Wagner, L.H., Shenmen, C.M., Schuler, G.D.,  
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,  
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Haieh, F.,  
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,  
Scapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,  
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,  
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,  
Abramson, R.D., Mullaby, S.J., Bosak, S.A., McEwan, P.J.,  
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,  
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,  
Vallalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,  
Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,  
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,  
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,  
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,  
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smal, D.E.,  
Scherer, A., Schein, J.E., Jones, S.J. and Marra, M.A.  
Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
12477932

2 (bases 1 to 1761)  
AUTHORS Strausberg, R.  
TITLE Direct Submission  
JOURNAL Submitted (03-DEC-2001) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
NTH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgabsa@mail.nih.gov](mailto:cgabsa@mail.nih.gov)  
Tissue Procurement: ATCC  
CNA Library Preparation: Life Technologies, Inc.  
CNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Sequencing Group at the Stanford Human Genome  
Center, Stanford University School of Medicine, Stanford, CA 94305  
Web site: <http://www-shgc.stanford.edu>  
Contact: (Dickson, Mark) [mcdepaxil.stanford.edu](mailto:mcdepaxil.stanford.edu)  
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,  
R. M.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAC Plate: 27 Row: h Column: 10  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA gi: 4506868.  
Location/Qualifiers

source	1. .1761		COMMENT	Protein Research Group; 1-7-22 Suehiro, Tsurumi, Yokohama, Kanagawa, 230-0045, Japan (E-mail:aktanaka@riken.jp, URL:http://protein.gsc.riken.jp/, Tel:81-45-503-9452, Fax:81-45-503-9450)		
	/organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="MGC:21265 IMAGE:4413387" /tissue_type="Liver, adenocarcinoma" /clone_lib="NIH MGC_90" /lab_host="DH10B" /note="Vector: pCMV-SPORT6"			This work was supported in part by the National Project on Protein Structural and Functional Analysis, Ministry of Education, Culture, Sports, Science and Technology of Japan. Sumio Sugano, Yutaka Suzuki Laboratory of Functional Genomics Department of Medical Genome Sciences Graduate School of Frontier Sciences The University of Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639 Japan email: ssugano@k.u-tokyo.ac.jp URL: http://www.k.u-tokyo.ac.jp/index.html.en.		
gene	1. .1761		FEATURES	Location/Qualifiers		
	/gene="SECTM1" /notes="synonym: K12" /db_xref="GeneID:6398" /db_xref="MIM:602602" 93. .839 /gene="SECTM1" /codon_start=1 /product="secreted and transmembrane 1 precursor" /protein_id="AAH17716.1" /db_xref="GI:17389346" /db_xref="GeneID:6398" /db_xref="MIM:602602" /translation="MOTCLPAPGHVSQALGTLFLFLAASLSAQNEGWDSPICTEGVVS VSWGENTVMSCNISNAFSHVNIKLRAHQESAIENEVAPGYFSRDGMQLQVQGGVAQL VIKGARDSHAGLYMHVLVGHORNNRQVTLVSGAEPQSDPTGFWPVPVAVTAVFILL VALVMPAWYRCRCQQORREKFFLLEPQMKFAALRAGAQQGLSRASAELETPDSEPTP RPLALVFKPSPLGALLELLSPQLFPFYAADP"			1. .1766 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="KAT02983" /cell_line="KATO III" /clone_lib="KAT" /note="cloning vector: pME18SFL3 this clone is also named as hss001002587" <92. .838 /note="Start codon is not identified." /codon_start=1 /evidence="not experimental" /product="secreted and transmembrane 1 precursor variant" /protein_id="BAD96788.1" /db_xref="GI:62897697" /translation="MOTCLPAPGHVSQALGTLFLFLAASLSAQNEGWDSPICTEGVVS VSWGENTVMSCNISNAFSHVNIKLRAHQESAIENEVAPGYFSRDGMQLQVQGGVAQL VIKGARDSHAGLYMHVLVGHORNNRQVTLVSGAEPQSDPTGFWPVPVAVTAVFILL VALVMPAWYRCRCQQORREKFFLLEPQMKFAALRAGAQQGLSRASAELETPDSEPTP RPLALVFKPSPLGALLELLSPQLFPFYAADP"		
CDS				CDS		
ORIGIN	Query Match 94.7%; Score 74.8; DB 8; Length 1761; Best Local Similarity 97.4%; Pred. No. 2.9e-11; Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;		ORIGIN	Query Match 94.7%; Score 74.8; DB 8; Length 1766; Best Local Similarity 97.4%; Pred. No. 2.9e-11; Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;		
	QY 1 CGGGAATCCCTTCAGTTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60       DB 1620 CGGGAATCCCTTCAGTTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1679 			QY 1 CGGGAATCCCTTCAGTTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60       DB 1619 CGGGAATCCCTTCAGTTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1678 		
RESULT 7	61 AACAAACCGACTCTAAA 78 		RESULT 8	CQ776411 2000 bp DNA linear PAT 11-MAR-2004		
	LOCUS AK223068 1766 bp mRNA linear PRI 26-APR-2005			Sequence 97 from Patent EP1394274.		
DEFINITION	Homo sapiens mRNA for secreted and transmembrane 1 precursor variant, clone: KAT02983.		DEFINITION	Homo sapiens (human)		
ACCESSION	AK223068		ACCESSION	CQ776411		
VERSION	AK223068.1 GI:62897696		VERSION	CQ776411.1 GI:45379801		
KEYWORDS	FLI CDNA; oligo capping.		KEYWORDS	Homo sapiens (human)		
SOURCE	Homo sapiens		SOURCE	Homo sapiens		
ORGANISM	Homo sapiens		ORGANISM	Homo sapiens		
REFERENCE	1		REFERENCE	1		
	Suzuki, Y., Yoshitomo-Nakagawa, K., Maruyama, K., Suyama, A. and Sugano, S.			Ohtani, N., Sugita, Y., Yamaya, M., Kubo, H., Nagai, H. and Izuahara, K.		
AUTHORS	Suzuki, Y., Sugano, S., Totoki, Y., Toyoda, A., Takeda, T., Sakaki, Y., Tanaka, A. and Yokoyama, S.		AUTHORS	Methods of testing for bronchial asthma or chronic obstructive pulmonary disease		
TITLE	Oligo-capping: a simple method to replace the cap structure of eukaryotic mRNAs with oligoribonucleotides		TITLE	Patent: EP 1394274-A 97 03-MAR-2004;		
JOURNAL	Gene 138 (1-2), 171-174 (1994)		JOURNAL	Genox Research, Inc. (JP)		
PUBLISHED	8125298		PUBLISHED	Location/Qualifiers		
REFERENCE	2		REFERENCE	1. .2000		
AUTHORS	Suzuki, Y., Yoshitomo-Nakagawa, K., Maruyama, K., Suyama, A. and Sugano, S.			/organism="Homo sapiens"		
TITLE	Construction and characterization of a full length-enriched and a 5'-end-enriched cDNA library			/mol_type="unassigned DNA"		
JOURNAL	Gene 200 (1-2), 149-156 (1997)			/db_xref="taxon:9606"		
PUBLISHED	9373149					
REFERENCE	3					
AUTHORS	(bases 1 to 1766)					
TITLE	Suzuki, Y., Sugano, S., Totoki, Y., Toyoda, A., Takeda, T., Sakaki, Y., Tanaka, A. and Yokoyama, S.					
JOURNAL	Direct Submission					
PUBLISHED	Submitted (22-APR-2005) Akiko Tanaka, RIKEN Yokohama Institute,					

```

ORIGIN
Query Match          94.7%; Score 74.8; DB 6; Length 2000;
Best Local Similarity 97.4%; Pred. No. 2.8e-11;
Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
    |||||
Db 1645 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1704
    |||||

Qy 61 AACAAACCGACTCTAAAA 78
    |||||
Db 1705 AACAAACCGACTCAGAAA 1722

RESULT 9
LOCUS      CQ861464          2000 bp      DNA      linear      PAT 10-SEP-2004
DEFINITION Sequence 97 from Patent WO2004072265.
ACCESSION  CQ861464
VERSION     CQ861464.1 GI:51982453
KEYWORDS   Homo sapiens (human)
SOURCE     Homo sapiens
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE  1 Burczynski, M., Twine, N., Dörner, A.J. and Trepicchio, W.L.
AUTHORS   METHODS FOR MONITORING DRUG ACTIVITIES IN VIVO / I
TITLE     Patent: WO 2004072265-A 97 26-AUG-2004;
JOURNAL   Wyeth (US); Burczynski, Michael E. (US); Twine, Natalie C. (US);
            Dörner, Andrew J. (US); Trepicchio, William L. (US)
FEATURES   Location/Qualifiers
            source          1..2000
                        /organism="Homo sapiens"
                        /mol_type="unassigned DNA"
                        /db_xref="taxon:9606"

ORIGIN
Query Match          94.7%; Score 74.8; DB 6; Length 2000;
Best Local Similarity 97.4%; Pred. No. 2.8e-11;
Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
    |||||
Db 1645 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1704
    |||||

Qy 61 AACAAACCGACTCTAAAA 78
    |||||
Db 1705 AACAAACCGACTCAGAAA 1722

RESULT 10
LOCUS      CQ981630          2000 bp      DNA      linear      PAT 25-JAN-2005
DEFINITION Sequence 485 from Patent EP1498424.
ACCESSION  CQ981630
VERSION     CQ981630.1 GI:58190920
KEYWORDS   Homo sapiens (human)
SOURCE     Homo sapiens
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE  1 Rosenthal, A., Hermann, K., Heiden, E., Pilarsky, C., Brummendorf, T.,
AUTHORS   Staub, E., Roepcke, S., Mennerich, D., Kinnemann, H. and Li, X.
TITLE     Human nucleic acid sequences from lung tumours
JOURNAL   Patent: EP 1498424-A 485 19-JAN-2005;
            Hinzmann, Bernd (DE); Hermann, Klaus (DE); Heiden, Esmeralda (DE);
            Rosenthal, Andre (DE)
FEATURES   Location/Qualifiers

```

```

source          1..2000
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match          94.7%; Score 74.8; DB 6; Length 2000;
Best Local Similarity 97.4%; Pred. No. 2.8e-11;
Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
    |||||
Db 1645 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1704
    |||||

Qy 61 AACAAACCGACTCTAAAA 78
    |||||
Db 1705 AACAAACCGACTCAGAAA 1722

RESULT 11
LOCUS      AR569106          2000 bp      DNA      linear      PAT 14-DEC-2004
DEFINITION Sequence 3 from patent US 6762030.
ACCESSION  AR569106
VERSION     AR569106.1 GI:56569141
KEYWORDS   Unknown.
SOURCE     Unknown.
ORGANISM   Unclassified.
REFERENCE  1 (bases 1 to 2000)
AUTHORS   Lyman, S.D. and Fanslow, W.C. III.
TITLE     Ligand for CD7 and methods for use thereof
JOURNAL   Patent: US 6762030-A 3 13-JUL-2004;
            Immunex Corporation; Seattle, WA,
            WOX;
FEATURES   Location/Qualifiers
            source          1..2000
                        /organism="unknown"
                        /mol_type="genomic DNA"

ORIGIN
Query Match          94.7%; Score 74.8; DB 6; Length 2000;
Best Local Similarity 97.4%; Pred. No. 2.8e-11;
Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
    |||||
Db 1645 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1704
    |||||

Qy 61 AACAAACCGACTCTAAAA 78
    |||||
Db 1705 AACAAACCGACTCAGAAA 1722

RESULT 12
LOCUS      AX055560          2000 bp      DNA      linear      PAT 13-JAN-2001
DEFINITION Sequence 3 from Patent WO0073333.
ACCESSION  AX055560
VERSION     AX055560.1 GI:12228778
KEYWORDS   Homo sapiens (human)
SOURCE     Homo sapiens
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE  1 Lyman, S.D. and Fanslow, W.C.
AUTHORS   Ligand for cd7, and methods for use thereof
TITLE     Patent: WO 0073333-A 3 07-DEC-2000;
JOURNAL   Immunex Corporation (US)
FEATURES   Location/Qualifiers
            source          1..2000

```

KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE  
AUTHORS  
TITLE  
1 (bases 1 to 2000)  
Slentz-Kesler, K.A., Hale, L.P. and Kaufman, R.E.  
Identification and characterization of K12 (SECTM1), a novel human  
gene that encodes a Golgi-associated protein with transmembrane and  
secreted isoforms  
Genomics 47 (3), 327-340 (1998)  
JOURNAL  
PUBMED  
9480746  
REFERENCE  
AUTHORS  
TITLE  
Slentz-Kesler, K.A. and Kaufman, R.E.  
Direct Submission  
JOURNAL  
DUMC, Durham, NC 27710, USA  
FEATURES  
Location/Qualifiers  
1..2000  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="17"  
/map="17q25.2-25.3"  
/cell\_type="K562 erythroleukemic"  
119..865  
/note="type Ia transmembrane protein"  
/codon\_start=1  
/product="K12 protein precursor"  
/protein\_id="AAC52044.1"  
/db\_xref="GI:2062391"  
/translation="MQTCPLAFPHGVHQAQVSLFVNEVAPGYSRDSQWQVQGGVAAQL  
VSGENTVMSCNISNAPSHVNIKLRAHQGESAIFFNEVAPGYSRDSQWQVQGGVAAQL  
VSGENTVMSCNISNAPSHVNIKLRAHQGESAIFFNEVAPGYSRDSQWQVQGGVAAQL  
VALVMPAWYRCRCSQQRREKFFLLEPQMKVAALRAGAQQGLSRASAEIWTDPSEPTP  
RPLALVFKPSPLGALLELLSPQLFPYAADP"  
119..202  
sig\_peptide  
203..862  
mat\_peptide  
551..623  
misc\_feature  
/note="encodes transmembrane domain"  
752..862  
misc\_feature  
/note="encodes proline-rich region"  
ORIGIN  
Query Match 94.7%; Score 74.8; DB 6; Length 2000;  
Best Local Similarity 97.4%; Pled. No. 2.8e-11;  
Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 CGGGAATCCCCCTTCAGTCTTTTGAAGAAAGTTCATGACTCGAATATCTGAAATGAAGAA 60  
Db 1645 CGGGAATCCCCCTTCAGTCTTTTGAAGAAAGTTCATGACTCGAATATCTGAAATGAAGAA 1704  
Qy 61 AACAAACCGACTCAAAA 78  
Db 1705 AACAAACCGACTCAAAA 1722  
RESULT 15  
AR071760  
LOCUS  
DEFINITION  
AR071760  
ACCESSION  
AR071760  
VERSION  
AR071760.1  
GI:7222648  
KEYWORDS  
SOURCE  
ORGANISM  
Unknown.  
Unclassified.  
REFERENCE  
AUTHORS  
TITLE  
Kaufman, R.E. and Slentz-Kesler, K.  
Gene product over expressed in cancer cells  
JOURNAL  
Patent: US 5912142-A 2 15-JUN-1999;

100  
119..865  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
119..865  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAC22100.1"  
/db\_xref="GI:12228779"  
/translation="MQTCPLAFPHGVHQAQVSLFVNEVAPGYSRDSQWQVQGGVAAQL  
VSGENTVMSCNISNAPSHVNIKLRAHQGESAIFFNEVAPGYSRDSQWQVQGGVAAQL  
VSGENTVMSCNISNAPSHVNIKLRAHQGESAIFFNEVAPGYSRDSQWQVQGGVAAQL  
VSGENTVMSCNISNAPSHVNIKLRAHQGESAIFFNEVAPGYSRDSQWQVQGGVAAQL  
VALVMPAWYRCRCSQQRREKFFLLEPQMKVAALRAGAQQGLSRASAEIWTDPSEPTP  
RPLALVFKPSPLGALLELLSPQLFPYAADP"  
94.7%; Score 74.8; DB 6; Length 2000;  
Query Match  
Best Local Similarity 97.4%; Pled. No. 2.8e-11;  
Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 CGGGAATCCCCCTTCAGTCTTTTGAAGAAAGTTCATGACTCGAATATCTGAAATGAAGAA 60  
Db 1645 CGGGAATCCCCCTTCAGTCTTTTGAAGAAAGTTCATGACTCGAATATCTGAAATGAAGAA 1704  
Qy 61 AACAAACCGACTCAAAA 78  
Db 1705 AACAAACCGACTCAAAA 1722  
RESULT 14  
HSU77643  
LOCUS  
DEFINITION  
HSU77643  
ACCESSION  
U77643  
VERSION  
U77643.1  
GI:2062390  
2000 bp mRNA  
linear  
PRI 10-MAR-1998  
Homo sapiens K12 protein precursor mRNA, complete cds.  
Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE  
AUTHORS  
TITLE  
Young, P.B., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,  
Horrikan, S., Soppet, D.R. and Weaver, Z.  
Cancer gene determination and therapeutic screening using signature  
gene sets  
Patent: WO 0194629-A 4264 13-DEC-2001;  
Avalon Pharmaceuticals (US)  
FEATURES  
Location/Qualifiers  
1..2000  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 94.7%; Score 74.8; DB 6; Length 2000;  
Best Local Similarity 97.4%; Pled. No. 2.8e-11;  
Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 CGGGAATCCCCCTTCAGTCTTTTGAAGAAAGTTCATGACTCGAATATCTGAAATGAAGAA 60  
Db 1645 CGGGAATCCCCCTTCAGTCTTTTGAAGAAAGTTCATGACTCGAATATCTGAAATGAAGAA 1704  
Qy 61 AACAAACCGACTCAAAA 78  
Db 1705 AACAAACCGACTCAAAA 1722

KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE  
AUTHORS  
TITLE  
1 (bases 1 to 2000)  
Slentz-Kesler, K.A., Hale, L.P. and Kaufman, R.E.  
Identification and characterization of K12 (SECTM1), a novel human  
gene that encodes a Golgi-associated protein with transmembrane and  
secreted isoforms  
Genomics 47 (3), 327-340 (1998)  
JOURNAL  
PUBMED  
9480746  
REFERENCE  
AUTHORS  
TITLE  
Slentz-Kesler, K.A. and Kaufman, R.E.  
Direct Submission  
JOURNAL  
DUMC, Durham, NC 27710, USA  
FEATURES  
Location/Qualifiers  
1..2000  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="17"  
/map="17q25.2-25.3"  
/cell\_type="K562 erythroleukemic"  
119..865  
/note="type Ia transmembrane protein"  
/codon\_start=1  
/product="K12 protein precursor"  
/protein\_id="AAC52044.1"  
/db\_xref="GI:2062391"  
/translation="MQTCPLAFPHGVHQAQVSLFVNEVAPGYSRDSQWQVQGGVAAQL  
VSGENTVMSCNISNAPSHVNIKLRAHQGESAIFFNEVAPGYSRDSQWQVQGGVAAQL  
VSGENTVMSCNISNAPSHVNIKLRAHQGESAIFFNEVAPGYSRDSQWQVQGGVAAQL  
VSGENTVMSCNISNAPSHVNIKLRAHQGESAIFFNEVAPGYSRDSQWQVQGGVAAQL  
VALVMPAWYRCRCSQQRREKFFLLEPQMKVAALRAGAQQGLSRASAEIWTDPSEPTP  
RPLALVFKPSPLGALLELLSPQLFPYAADP"  
119..202  
sig\_peptide  
203..862  
mat\_peptide  
551..623  
misc\_feature  
/note="encodes transmembrane domain"  
752..862  
misc\_feature  
/note="encodes proline-rich region"  
ORIGIN  
Query Match 94.7%; Score 74.8; DB 8; Length 2000;  
Best Local Similarity 97.4%; Pled. No. 2.8e-11;  
Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 CGGGAATCCCCCTTCAGTCTTTTGAAGAAAGTTCATGACTCGAATATCTGAAATGAAGAA 60  
Db 1645 CGGGAATCCCCCTTCAGTCTTTTGAAGAAAGTTCATGACTCGAATATCTGAAATGAAGAA 1704  
Qy 61 AACAAACCGACTCAAAA 78  
Db 1705 AACAAACCGACTCAAAA 1722  
RESULT 15  
AR071760  
LOCUS  
DEFINITION  
AR071760  
ACCESSION  
AR071760  
VERSION  
AR071760.1  
GI:7222648  
KEYWORDS  
SOURCE  
ORGANISM  
Unknown.  
Unclassified.  
REFERENCE  
AUTHORS  
TITLE  
Kaufman, R.E. and Slentz-Kesler, K.  
Gene product over expressed in cancer cells  
JOURNAL  
Patent: US 5912142-A 2 15-JUN-1999;

2180 bp  
DNA  
linear  
PAT 18-FEB-2000  
Sequence 2 from patent US 5912142.  
AR071760  
AR071760.1  
GI:7222648

FEATURES  
source Location/Qualifiers  
1..2180  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN

Query Match 94.7%; Score 74.8; DB 6; Length 2180;  
Best Local Similarity 97.4%; Pred. No. 2.7e-11;  
Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60  
|||  
Db 1841 CGGGAATCCCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1900  
|||

Qy 61 AACAAACCGACTCTAAAA 78  
|||  
Db 1901 AACAAACCGACTCAGAAA 1918  
|||

Search completed: December 7, 2005, 04:21:56  
Job time : 703.666 secs

*This Page Blank (uspr,*



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 21:43:57 ; Search time 3355.79 Seconds  
(without alignments)  
4099.008 Million cell updates/sec

Title: US-09-980-046B-2  
Perfect score: 294  
Sequence: 1 cggccttaaggttcctctga.....aagagagtaattccccaaaa 294

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 41078125 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_est4:\*  
5: gb\_est5:\*  
6: gb\_est6:\*  
7: gb\_est7:\*  
8: gb\_est8:\*  
9: gb\_est9:\*  
10: gb\_est10:\*  
11: gb\_est11:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	289.2	98.4	561	BF195984	7088e02.x
C 2	275.6	93.7	599	AW082845	AW082845 xc03h03.x
C 3	228.4	77.7	540	AQ167893	AQ167893 HS 2224 B
C 4	50	17.0	939	DU045370	DU045370 153083 To
C 5	49.6	16.9	807	CX948891	CX948891 UMC-bcl 0
C 6	49.4	16.8	909	CZ514707	CZ514707 GWM2-9TL2
C 7	49.4	16.8	928	CNS00DKY	AL071865 Drosophil
C 8	49.4	16.8	1101	CNS00EVL	AL069706 Drosophil
C 9	49.2	16.7	1101	CNS01450	AL103740 Drosophil
C 10	49	16.7	987	CNS014PQ	AL104456 Drosophil
C 11	48.2	16.4	897	CNS015FK	AL105386 Drosophil
C 12	48	16.3	979	CNS016LW	AL106190 Drosophil
C 13	48	16.3	1101	CNS016LW	AL108896 Drosophil
C 14	47.8	16.3	1101	CNS01219	AL101595 Drosophil
C 15	47.2	16.1	572	CX572895	CX572895 TTE000224
C 16	47	16.0	1432	AG381846	AG381846 Mus muscu
C 17	46.8	15.9	1201	CNS0107Y	AL098632 Drosophil
C 18	46.2	15.7	836	DN078399	DN078399 JGI_CASD1
C 19	46.2	15.7	836	CNS00DKG2	AL077933 Drosophil
C 20	46.2	15.7	964	BX426198	BX426198 BX426198
C 21	45.8	15.6	688	AG167069	AG167069 Pan trogl
C 22	45.8	15.6	811	AL514901	AL514901 AL514901

23	45.8	15.6	813	9	BZ512051	BZ512051 BOMQI3TF
C 24	45.6	15.5	479	8	DN344837	DN344837 LI83551-0
C 25	45.6	15.5	652	10	CZ143061	CZ143061 OA BBA004
C 26	45.6	15.5	771	9	AV758683	AV758683 RP43-121h
27	45.4	15.4	1101	10	CNS002AL	AL097431 Drosophil
28	45.2	15.4	713	9	BZ043743	BZ043743 lj30f06.
29	45.2	15.4	928	10	CNS00DKY	AL071865 Drosophil
C 30	45	15.3	850	10	CZ543128	CZ543128 SRAA-aad4
C 31	45	15.3	894	9	AQ621876	AQ621876 HS 3107 B
C 32	44.8	15.2	759	11	CNS06QXV	AL411257 T7 end of
C 33	44.6	15.2	285	3	BM893703	BM893703 ij28f12.x
C 34	44.6	15.2	492	1	AU053286	AU053286 AU053286
C 35	44.4	15.1	481	11	DE137325	DE137325 Oryzias 1
36	44.4	15.1	744	7	CJ344067	CJ344067 CJ344067
37	44.4	15.1	951	11	CNS07BHQ	AL437892 T7 end of
C 38	44.4	15.1	992	10	CL144176	CL144176 ISB1-144B
C 39	44.4	15.1	1005	11	CNS07BO3	AL438121 T7 end of
40	44.4	15.1	1043	11	CNS07B49	AL437407 T3 end of
C 41	44.4	15.1	1298	10	AG360927	AG360927 Mus muscu
C 42	44.2	15.0	806	10	CNS008OW	AL052249 Drosophil
C 43	44.2	15.0	920	9	AZ671046	AZ671046 ENTKE23TR
C 44	44	15.0	624	8	CX573684	CX573684 TTE000312
45	44	15.0	747	10	AG364182	AG364182 Mus muscu

ALIGNMENTS

RESULT 1  
BF195984/c  
LOCUS

DEFINITION

7088e02.x1 NCI CGAP Kid11 Homo sapiens CDNA clone IMAGE:3643467 3', mRNA similar to SW:STB1\_BOVIN Q28208 SYNTAXIN BINDING PROTEIN 1 ; , mRNA sequence.

ACCESSION

BF195984

VERSION

BF195984.1

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 561)

NCI-CCGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: [cgapsb@mail.nih.gov](mailto:cgapsb@mail.nih.gov)

Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: M. Bento Soares, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CCGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL, send email to: [info@image.llnl.gov](mailto:info@image.llnl.gov)

High quality sequence stop: 490.

Location/Qualifiers

1..561

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:3643467"

/lab\_host="DH10B"

/clone\_lib="NCI\_CGAP Kid11"

/note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI;

Plasmid DNA from the normalized library NCI\_CGAP Kid3 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library

FEATURES  
source

561 bp mRNA linear  
7088e02.x1 NCI CGAP Kid11 Homo sapiens CDNA clone IMAGE:3643467 3', mRNA similar to SW:STB1\_BOVIN Q28208 SYNTAXIN BINDING PROTEIN 1 ; , mRNA sequence.

BF195984

BF195984.1

EST

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 561)

NCI-CCGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: [cgapsb@mail.nih.gov](mailto:cgapsb@mail.nih.gov)

Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: M. Bento Soares, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CCGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL, send email to: [info@image.llnl.gov](mailto:info@image.llnl.gov)

High quality sequence stop: 490.

Location/Qualifiers

1..561

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:3643467"

/lab\_host="DH10B"

/clone\_lib="NCI\_CGAP Kid11"

/note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI;

Plasmid DNA from the normalized library NCI\_CGAP Kid3 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library

(cloneIDs 1322376-1323911, 1456007-1456775, and 1500552-1502855). Subtraction by Bento Soares and M. Fatima Bonaldo.

```

ORIGIN
Query Match      98.4%; Score 289.2; DB 2; Length 561;
Best Local Similarity 99.0%; Pred. No. 2e-53;
Matches 291; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CGGCCTTAAGGTTCTCTGACAGTGTCTCAAGAAGATCAGGATTTGAGCATCTCTTACT 60
   |||
Db 294 CGGCCTTAAGGTTCTCTGACAGTGTCTCAAGAAGATCAGGATTTGAGCATCTCTTACT 235
   |||

QY 61 TCAAAATTAATCTTCCAGTGTGTTTGTATCTCTCAATGTTATAGTGAATAATAATA 120
   |||
Db 234 TCAAAATTAATCTTCCAGTGTGTTTGTATCTCTCAATGTTATAGTGAATAATAATA 175
   |||

QY 121 CATGCAGAAAAAGCATATAATCAACTATTGAATAACGAAATGTGAACCTCTATGTA 180
   |||
Db 174 CATGCAGAAAAAGCATATAATCAACTATTGAATAACGAAATGTGAACCTCTATGTA 115
   |||

QY 181 GCAACTCAGGCTGCTACCGTTTAAATTTCTCTTAATCTTTTAAATAACAAATAC 240
   |||
Db 114 GCAACTCAGGCTGCTACCGTTTAAATTTCTCTTAATCTTTTAAATAACAAATAC 55
   |||

QY 241 TTACTAAAGATGTTTCAGAGTAAGATAAATACAAAGAGTAAATCCCAAAA 294
   |||
Db 54 TTACTAAAGATGTTTCAGAGTAAGATAAATACAAAGAGTAAATCTCCAAA 1
   |||

```

```

RESULT 2
AW082845/c
LOCUS
DEFINITION
  xc03h03.x1 NCI CGAP C021 Homo sapiens cDNA clone IMAGE:2583221 3,
  similar to SW:STX1_BOVIN Q28208 SYNTAXIN BINDING PROTEIN 1 ; mRNA
  sequence.
ACCESSION
  AW082845
VERSION
  AW082845.1 GI:6037997
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
  Homiidae; Homo.
REFERENCE
  1 (bases 1 to 599)
  NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
  Unpublished (1997)
  Contact: Robert Strausberg, Ph.D.
  Email: cgaps-remail.nih.gov
  Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
  R. Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: Life
  Technologies, Inc. cDNA Library Arrayed by: Christa Prange, The
  I.M.A.G.E. Consortium DNA Sequencing by: Washington University
  Genome Sequencing Center
  Clone distribution: NCI-CGAP clone distribution information can be
  found through the I.M.A.G.E. Consortium/LLNL at:
  www-bio.llnl.gov/bbrp/image/image.html
  Seq primer: -40UP from Gibco
  High quality sequence stop: 417.
  Location/Qualifiers
    1..599
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="IMAGE:2583221"
      /tissue_type="moderately differentiated adenocarcinoma"
      /lab_host="DH10B"
      /clone_lib="NCI CGAP C021"
      /notes="Organ: colon; Vector: pCMV-SPORT6; Site 1: SalI;
      Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
      Normalized to Cot >500. Average insert size 1.04kb.

```

Normalized version of NCI\_CGAP\_Col8. Library constructed by Life Technologies."

```

ORIGIN
Query Match      93.7%; Score 275.6; DB 1; Length 599;
Best Local Similarity 98.3%; Pred. No. 2e-50;
Matches 289; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 1 CGGCCTTAAGGTTCTCTGACAGTGTCTCAAGAAGATCAGGATTTGAGCATCTCTTACT 60
   |||
Db 302 CGGCCTTAAGGTTCTCTGACAGTGTCTCAAGAAGATCAGGATTTGAGCATCTCTTACC 244
   |||

QY 61 TCAAAATTAATCTTCCAGTGTGTTTGTATCTCTCAATGTTATAGTGAATAATAATA 120
   |||
Db 243 TCAAAATTAATCTTCCAGTGTGTTTGTATCTCTCAATGTTATAGTGAATAATAATA 184
   |||

QY 121 CATGCAGAAAAAGCATATAATCAACTATTGAATAACGAAATGTGAACCTCTATGTA 180
   |||
Db 183 CATGCAGAAAAAGCATATAATCAACTATTGAATAACGAAATGTGAACCTCTATGTA 124
   |||

QY 181 GCAACTCAGGCTGCTACCGTTTAAATTTCTCTTAATCTTTTAAATAACAAATAC 240
   |||
Db 123 GCAACTCAGGCTGCTACCGTTTAAATTTCTCTTAATCTTTTAAATAACAAATAC 64
   |||

QY 241 TTACTAAAGATGTTTCAGAGTAAGATAAATACAAAGAGTAAATCCCAAAA 294
   |||
Db 63 TTACTAAAGATGTTTCAGAGTAAGATAAATACAAAGAGTAAATCTCCAAA 10
   |||

```

```

RESULT 3
AW0167893
LOCUS
DEFINITION
  HS_2224_B2_G09_MR CIT Approved Human Genomic Sperm Library D Homo
  sapiens genomic clone Plate=2224 Col=18 Row=N, genomic survey
  sequence.
ACCESSION
  AW0167893
VERSION
  AW0167893.1 GI:3561608
KEYWORDS
  GSS.
SOURCE
  Homo sapiens (human)
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
  Homiidae; Homo.
REFERENCE
  1 (bases 1 to 540)
  Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., and
  Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
  Hood,L.
  Sequence-tagged connectors: A sequence approach to mapping and
  scanning the human genome
  Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
  10449764
  Contact: Mahairas GG, Wallace JC, Hood L
  High Throughput Sequencing Center
  University of Washington
  401 Queen Anne Avenue North, Seattle, WA 98109, USA
  Tel: (206) 616-3618
  Fax: (206) 616-3887
  Email: jwallace@u.washington.edu
  Sequence Tagged Connector
  Plate: 2224 row: N column: 18
  Class: BAC ends
  High quality sequence stop: 540.
  Location/Qualifiers
    1..540
      /organism="Homo sapiens"
      /mol_type="genomic DNA"
      /db_xref="taxon:9606"
      /clone="Plate=2224 Col=18 Row=N"
      /sex="male"
      /clone_lib="CIT Approved Human Genomic Sperm Library D"
      /notes="Organ: sperm; Vector: pBelBAC11; BAC Clones in
      E-Coli DH10B"

```

```

Query Match      77.7%; Score 228.4; DB 9; Length 540;
Best Local Similarity 93.7%; Pred. No. 4.8e-40;
Matches 238; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 1 CGGCCTTAAGTTCTCTGACAGTGTCTCAAGAAGATCAGGATTTGAGCATCTCTTACT 60
Db 259 CGGCCTTAAGTTCTCTGACAGTGTCTCAAGAAGATCAGGATTTGAGCATCTCTTACT 318

Qy 61 TCAATATTATTTCCAGTGGTTTGTATCTCTCAATGTTATAGTGAATATAATA 120
Db 319 TCAATATTATTTCCAGTGGTTTGTATCTCTCAATGTTATAGTGAATATAATA 378

Qy 121 CATGCAGAAAAAGCATATAAACAATATTGAATAACGAAATGTGAACCTCTATGTAAT 180
Db 379 CATGGCACAAGAGATTAATCAACTATTGAATACGAAATGTGAACCTCTATGTAAT 438

Qy 181 GCAACTCAGGCTGCCTACCGTTTTTAATTTCTCTCTAATCTTTAAACAAATACTATCT 240
Db 439 GCAACTCAGGCTGCCTACCGTTTTTAATTTCTCTCTAATCTTTAAACAAATACTATCT 498

Qy 241 TTACTTAAGATGTT 254
Db 499 CTACTTAGATGTT 512

RESULT 4
DU045370/c
LOCUS
DEFINITION DU045370 939 bp DNA linear GSS 12-AUG-2005
clone LE_HBA0025020 5, genomic survey sequence.
ACCESSION DU045370.1 GI:72471206
VERSION
KEYWORDS
SOURCE
ORGANISM Lycopersicon esculentum (Solanum lycopersicum)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE 1 (bases 1 to 939)
AUTHORS Mueller,L.A., Buels,R.M., Wang,Y., Tankley,S.D., Giovannoni,J.J.,
Van Eck,J. and Stack,S.
BAC end sequencing from three Solanum lycopersicon libraries
TITLE
JOURNAL
COMMENT Other_GSSs: 153082
Contact: Lukas Mueller
Tanksley Lab, Dept. of Plant Breeding
Cornell University
251 Emerson Hall, Ithaca, NY 14853, USA
Tel: 607-255-6557
Fax: 607-255-6683
Email: sgn-feedback@gn.cornell.edu
Insert Length: 104809 Std Error: 0.00
Plate: 25 row: 0 column: 20
Seq primer: T7
Classes: BAC ends
High quality sequence start: 51
High quality sequence stop: 931.
Location/Qualifiers
1..939
/organism="Lycopersicon esculentum"
/mol_type="genomic DNA"
/cultivar="Heinz 1706"
/db_xref="taxon:4081"
/clone="LE_HBA0025020"
/lab_host="E. coli"
/clone_lib="Tomato HindIII BAC Library"
/note="Vector: pBelobAC11; Site_1: HindIII"

ORIGIN
Query Match      17.0%; Score 50; DB 10; Length 939;
Best Local Similarity 53.6%; Pred. No. 0.83;
Matches 104; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

Qy 100 TTATAGTGAATATAATAATACATGACGAAAAAGCATATAATCAACTATTGAATAACGAA 159
Db 848 TAAATTTAGGGAATATAGAAAAAATAATAATAATTAATCACTCCAAATTTGACAAATAAA 789

Qy 160 AATGTGAACCTCTATGTAATCTGCAACTCAGGCTGCCTACCGTTTTTAATTTCTCTCTAAT 219
Db 788 AATAGTAAACCTTAAATCTACCTTTTTTATATATATTTTTTAAATTTTAAATTTTTCGATAAT 729

Qy 220 CTTTAAAAACAAATATCTATCTTTACTAAAGATGTTTCAGAGATGAATAATAACAAAGAG 279
Db 728 CTCAAAAATGAAATAATGTTAATAAAAAATATTTAAAAAAATTTAAAAAAGATATA 669

Qy 280 AGTAAATTTCCCAAAA 293
Db 668 TATAACCCCAAAA 655

CX948891 807 bp mRNA linear EST 08-FEB-2005
LOCUS
DEFINITION UMC-bcl_OA02-021-b09 Day 14 CL +8h prostaglandin F2-alpha bcl Bos
taurus cDNA 3', mRNA sequence.
ACCESSION CX948891
VERSION
KEYWORDS EST.
SOURCE
ORGANISM Bos taurus (cow)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Rutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
Pecora; Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 807)
AUTHORS Prather,R.S., Antoniou,E., Garverick,H.A., Green,J.A., Lucy,M.C.,
Roberts,R.M., Smith,M.P. and Youngquist,R.S.
USDA Grant MRI-2002-03476: Bovine ESTs: Focus on Female
Reproduction
JOURNAL
COMMENT Contact: DNA Core Facility (Bovine Project)
Animal Science - RS Prather
University of Missouri-Columbia
M16 Medical Sciences Bldg., Columbia, MO 65212, USA
Tel: (573)882-0428
Fax: (573)884-5552
Email: bovine@net.missouri.edu
POLYA=Yes.
Location/Qualifiers
1..807
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/clone_lib="bcl"
/note="Funding: The production of ESTs submitted in this
project was funded by USDA Grant MRI-2002-03476 entitled
'Bovine ESTs: Focus on Female Reproduction' to RS
Prather(Primary Investigator), E Antoniou, HA Garverick,
JA Green, MC Lucy, RM Roberts, MP Smith and RS Youngquist.
Genetic Source: Heifers for the project were purchased
from Circle A Ranch, Iberia, MO
(http://www.circlearanch.com/home.html). These heifers,
while not registered have known Angus pedigrees going back
at least 4 generations. Samples collected: The samples
consisted of the following: germinal vesicle-stage
oocytes; in vitro derived embryos (2-cell, morula,
blastocyst and nuclear transfer blastocyst); in vivo
blastocysts and conceptuses (days 8, 14, 16 and 18);
corpora lutea (days 3, 5, 8, 14, 16, 18 and 35); ovarian
follicles (days 0, non-recruited, recruited, early
selected and preovulatory); oviduct (days 0, 3 and 5);
endometrium (days 5, 8, 14, 16, 18 and 35); and
placenta/embryo from day 35 conceptuses. Expanded
descriptions of how the tissues were collected can be
found at the following URL:
http://genome.rnet.missouri.edu/Bovine/Methods.html.

```

Library construction (Standard Protocol): All procedures have been described in detail elsewhere (Soares et al., 1994; Bonaldo et al., 1996; Jiang et al., 2001). Total cellular RNA from each sample was isolated by using STAR-60 reagent (Tel-Test, Friendswood, TX) and the poly(A)+ RNA was obtained by two rounds of purification with the Oligotex mRNA isolation kit (Qiagen) according to the manufacturer's instructions. The libraries were constructed essentially as described by the manufacturer's instructions provided with the SuperScript Plasmid System (Invitrogen, cat. no. 18248-013). Briefly, 1mcg of poly(A)+ RNA was annealed at c37 degrees with 10mcg of NotI-tag-dT18 oligonucleotide (GCTGCTCGGCGC-tag-T18) and reverse transcribed at c37 degrees with SuperScript II (Invitrogen) reverse transcriptase (Jiang et al., 2001). The 'tag' represents a tissue/stage-specific ten-base sequence identifier (<http://genome.uiowa.edu/pubsoft/software.html>) present in the oligonucleotide used to prime first-strand synthesis. Second strand synthesis was performed with T4 DNA polymerase in the presence of DNA ligase and RNase H. After second strand synthesis, the double-stranded cDNAs were ligated to SalI adapters (Invitrogen-Life Technologies) and digested with NotI. The cDNAs were size selected by passage through cDNA size fractionation columns (Invitrogen-Life technologies). The cDNAs derived from each developmental stage of a particular tissue were mixed on an equimolar basis and ligated directionally into the NotI and SalI sites of the pCMV-SPORT6 vector (Invitrogen). After ligation of the inserts, the plasmids were electroporated into DH10B bacteria. Library construction (PCR Protocol): The amount of mRNA that was recovered from oocytes and embryos was quite limiting and was not sufficient for library production with the standard protocol. Therefore, PCR-based protocol was utilized for producing libraries from sources in which the amount of extracted mRNA was small (oocytes and embryos). Poly-A RNA was isolated by using the MicroPoly(A) Pure kit from Ambion (cat. # 1918). The mRNA was reverse transcribed with a NotI-tag-dT18 oligonucleotide and a SMART oligonucleotide (Clontech) modified to contain a SalI site to generate full-length cDNA with a sequence complementary to the SMART oligonucleotide. Sequences within the SMART and dT oligonucleotides were used as primers to amplify the cDNAs by PCR with pfu turbo polymerase (Stratagene). The resulting PCR products were purified, digested with NotI and SalI and size fractionated by using Chroma Spin-1000 columns (Clontech). Purified cDNA from each PCR reaction was quantitated and mixed on an equimolar basis for ligation into the pCMV-SPORT6 vector. Preliminary Library Characterization: Randomly chosen clones from each library were analyzed by restriction digestion to determine average insert size (96 clones) and by sequencing (-4 96-well plates) to confirm library quality (e.g. the presence of short polyA+ tails, genomic DNA contamination (must be <1%), ribosomal RNA clones (must be <1%), etc.) and to provide a sequence database representing the predominant clones in each library. The clones were sequenced at the University of Missouri-Columbia DNA Core Facility. After production of the libraries, equal numbers of recombinants from each library were pooled to produce a single mixed library (mega-library) for more extensive sequencing. Bioinformatics work was performed by GK Springer's bioinformatics group in Computer Science at the University of Missouri-Columbia. Clone Requests: Requests for clones should be made to the Director of the University of Missouri DNA Core facility at: [bovine@net.missouri.edu](mailto:bovine@net.missouri.edu). Bonaldo MF, Lennon G, Soares MB.. Normalization and Subtraction: Two approaches to facilitate gene discovery. Genome Res, 1996; 6:791-806. Jiang H, Bivens NJ, Ries JE, Whitworth KM, Green JA, Forrester LJ, Springer GK, Didion BA, Mathialagan N, Prather RS, Lucy MC (2001) Constructing

cDNA libraries with fewer clones that contain long poly(dA) tails. Biotechniques 31:38-42. Soares MB, MF Bonaldo, P Jelene, L Su, L Lawton, A Efstathiadis. 1994. Construction and characterization of a normalized cDNA library. Proc Natl Acad Sci, 91:9228-9232. TAG ISSUE=Day 14 CL +8h prostaglandin F2-alpha TAG\_SEQ=GAAGTGCTCC"

ORIGIN

Query Match	16.9%;	Score 49.6;	DB 8;	Length 807;
Best Local Similarity	50.9%;	Pred. No. 1;		
Matches	118;	Conservative	0;	Mismatches 114; Indels 0; Gaps 0;

QY 56 TTACTTCAATATTCTTCAGAGTGGTTTGGTTATCTTCTCAATGTTATAGTGAATATAA 115

Db 421 TTTCTCGATAAAATTTTTTTTTTGGCATTTTTTAAAGCTGTGCATTTACTTTCAAA 362

QY 116 TAATACATGCAGAAAAAGCATAAATCAACTATTGAATAACGAAAAATGTGAACCTCTATG 175

Db 361 TTTTAAGTCATGAAATATATATGCATTATATATGTAATAAGTTACATGAGAATTTCCCTG 302

QY 176 TAACTGCAACTCAGGCTGCCTACCGTTTAAATTTCTCTTAATCTTTTAAAAACAATAC 235

Db 301 TAACTACTGCTCAGAAATTTAAGATGTTTATATATCTTAAAAATATTTTTCAGAGCTAACTTA 242

QY 236 TATCTTTACTAAGATGTTTTCAGAGTAAGATAAAATACAAAGAGAGTAAATTC 287

Db 241 TTTATATGTTTTTACAGTTTCAGAGAAAAGTACAAGCAAGTTGTAAATTC 190

RESULT 6

CZ514707

LOCUS

DEFINITION

GW22-91L21a.g1 GW2 Glycine max genomic, genomic survey sequence.

ACCESSION

CZ514707

VERSION

CZ514707.1 GI:63239622

KEYWORDS

GSS.

SOURCE

Glycine max (soybean)

ORGANISM

Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.

REFERENCE

1 (bases 1 to 909)

Numberg,A., Bedell,J.A., Citek,R.W., Robbins,D., McMenamy,J., Peterson,S., Jones,J., Fries,J., Budiman,M.A., Nguyen,H., Clifton,S.W. and Stacey,G.

BAC end sequences from Glycine max Williams 82 cultivar genome

Unpublished (2005)

CONTACT: Richard K Wilson

Genome Sequencing Center

Washington University School of Medicine

Email: [submissions@watson.wustl.edu](mailto:submissions@watson.wustl.edu)

Class: BAC ends.

FEATURES

Location/Qualifiers

1..909

/organism="Glycine max"

/mol\_type="genomic DNA"

/db\_xref="taxon:3847"

/clone\_lib="GW2"

/notes="Glycine max BAC library"

ORIGIN

Query Match	16.8%;	Score 49.4;	DB 10;	Length 909;
Best Local Similarity	52.7%;	Pred. No. 1.1;		
Matches	107;	Conservative	0;	Mismatches 96; Indels 0; Gaps 0;

QY 35 AGATCAGGATTTGAGCATCTTACTTCAAAATTTCTTCCAGTGGTTTGTATCTTCT 94

Db 50 AAACAGAAAGTGCTGCTTCTCACTTTTAAAAAACCCCTTTAGTCTGCTTCTTCTTCTTT 109

QY 95 CAATGTTTAGTGAATATAATAATACATGCAGAAAAAGCATAAATCAACTATTGTAATA 154

Db 110 GAGTGGGTGGTGTGTTTGGAGAAGAGATAAATAAACAACAGAAATAGATAATTAATA 169

Qy 155 ACAGAAATGTCGAACCTCTATGTAAGTCAAGCTCAGGCTGCTACCGTGTCTCTCTCTCT 214

Db 170 ATGTAAATTTATATTTTATTTTAAATTTTAAATTTTATCTATTTTATTTTCTTT 229

Qy 215 CTAATCTTTTAAACAAATACTA 237

Db 230 CAACCTTTTAAACTATAAATA 252

RESULT 7

CNS00DKY/c

LOCUS

DEFINITION

Drosophila melanogaster genome survey sequence T7 end of BAC #

BACKR27A24 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

AL071865

AL071865.1 GI:4948170

GSS.

ORGANISM

Drosophila melanogaster (fruit fly)

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE

1 (bases 1 to 928)

Genoscope.

Direct Submission

TITLE

Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :

JOURNAL

BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr

- Web : www.genoscope.cns.fr)

COMMENT

Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).

FEATURES

source

Location/Qualifiers

1..928

/organism="Drosophila melanogaster"

/mol\_type="genomic DNA"

/db\_xref="taxon:7227"

/clone="BACR27A24"

/clone\_lib="RPCI-98"

/note="end : T7"

ORIGIN

Query Match 16.8%; Score 49.4; DB 10; Length 928;

Best Local Similarity 31.8%; Pred. No. 1.1;

Matches 92; Conservative 73; Mismatches 122; Indels 2; Gaps 1;

Qy 8 AAGGTTCCTCTGACAGTGTCTCAAGAAGATCAGGATTTGAGCATCTCTTACTTCAAAAT 67

Db 862 AAATTCMAAAMAMAMMMCMVMMMMMMMMMMMMMMMMMMMMMMHHHHHHHHTTTTTW 803

Qy 68 ATTCCTCCAGTGGTTTGTGTTATCTTCTCAATGTTATAGTGAATATAATAATACATGAG 127

Db 802 TTTTCTTTTCTTTTMM 743

Qy 128 AAAAAGCAATAATCAACTATTGAATACGAAATGTGAACCTCTATGTAAGTCAATC 187

Db 742 WATATTTTWTATWATAAAAAAATAAATAAATTTTCTTTTATTTTATTTTATTTTAAAA 683

Qy 188 AGGCTGCCTACCGTT--TTTATTTCTCTCTAATCTTTTAAAAACAATACTATCTTTACT 245

Db 682 TTTTATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 623

Qy 246 AAAGATGTTTTCAGAGTAAGATAAATAACAAGAGAGTAATTTCCCAAAA 294

Db 622 ATTTTWTATTWATATAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTAA 574

RESULT 8

CNS00EVL

LOCUS

DEFINITION

Drosophila melanogaster genome survey sequence T7 end of BAC:

BACR29B23 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

AL069706

AL069706.1 GI:4949849

GSS.

ORGANISM

Drosophila melanogaster (fruit fly)

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE

1 (bases 1 to 1101)

Genoscope.

Direct Submission

TITLE

Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :

JOURNAL

BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr

- Web : www.genoscope.cns.fr)

COMMENT

Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).

FEATURES

source

Location/Qualifiers

1..1101

/organism="Drosophila melanogaster"

/mol\_type="genomic DNA"

/db\_xref="taxon:7227"

/clone="BACR29B23"

/clone\_lib="RPCI-98"

/note="end : T7"

ORIGIN

Query Match 16.8%; Score 49.4; DB 10; Length 1101;

Best Local Similarity 37.9%; Pred. No. 1.1;

Matches 92; Conservative 40; Mismatches 111; Indels 0; Gaps 0;

Qy 52 TCTCTTACTTCAAAATTTCTTCCAGTGGTGTTCCTTCTTCTTCTCAATGTTATAGTGAAT 111

Db 633 TTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTA 692

Qy 112 ATAATAATACATGCAGAAAAAAGCATAAATCAACTATTGAAATAACGAAATGTGAACCTC 171

Db 693 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 752

Qy 172 TATGTAAGTCAACTCAGGCTGCCTACCGTGTCTTCTTCTTCTTCTTCTTCTTCTTCTT 231

Db 753 AATATWATATATATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTAA 812

Qy 232 ATACTATCTTTTACTTAAAGATGTTTTCAGATGAAGATAAATAACAAGAGAGTAATTTCC 291

Db 813 ATAWATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 872

```

QY      292 AAA 294
      :||
Db      873 WAA 875

RESULT 9
CNS0145U/c
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence T7 end of BAC
BACN11016 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION
VERSION AL103740.1 GI:5615351
KEYWORDS
SOURCE
ORGANISM
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1101)
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billaud at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelOBAC11.
FEATURES
source
Location/Qualifiers
1..1101
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACN11016"
/clone_lib="DrosBAC"
/plasmid="pBelOBAC11"
/note="end : T7"

ORIGIN
Query Match 16.7%; Score 49.2; DB 10; Length 1101;
Best Local Similarity 39.6%; Pred. No. 1.2;
Matches 93; Conservative 11; Mismatches 131; Indels 0; Gaps 0;

QY      60 TTCAAAATTATTCTCCAGTGGTTTGTGTTATCTCTCAATGTTATAGTGAATAATAAT 119
Db      368 TTTTNNNTNTTTTNTTNNNTNNNTNNNTNNNTNNNTNNNTNNNTNNNTNNNTNNNTNN 309

QY      120 ACATGCGAGAAAAAGCATAAATCAACTATTGAATAACGAAAAATGTGAACCTCTATGTAAC 179
Db      308 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 249

QY      180 TGCACCTCAGGCTGCCTACCGCTTTTAAATTTCTCTCTCTCTCTCTCTCTCTCTCTCT 239
Db      248 WTAATTATAATTTNNNNNTNTNTNNNTNTNTNNNTNTNTNNNTNNNTNNNTNNNTNN 189

QY      240 TTTTACTAAGATGTTTCAGAGTAAGATAAATAACAGAGAGTAAATCCCAAAA 294
Db      188 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 134

RESULT 10
CNS014PQ/c
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence SP6 end of BAC
BACN12P22 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION
VERSION AL104456
AL104456.1 GI:5616067

QY      64 AATTATTTCTCCAGTGGTTTGTGTTATCTCTCAATGTTATAGTGAATAATAATAAT 123
Db      913 WWWWWHTTTHWWTTTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWT 854

QY      124 GCAGAAAAAGCATAAATCAACTATTGAATAACGAAAAATGTGAACCTCTCTCTCTCTCT 183
Db      853 TTTTATAWAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 794

QY      184 ACTCAGGCTGCCTACCGCTTTTAAATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 243
Db      793 TAATTAATAWAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 734

QY      244 CTAAAGATGTTTCAGAGTAAGATAAATAACAAAGAGAGTAAT 285
Db      733 TTTWTTTTWTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 692

RESULT 11
CNS015FK
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence SP6 end of BAC
BACN13G14 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION
VERSION AL105386
AL105386.1 GI:5617400
KEYWORDS
SOURCE
ORGANISM
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 897)
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billaud at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelOBAC11.
FEATURES
source
Location/Qualifiers
1..987
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACN12P22"
/clone_lib="DrosBAC"
/plasmid="pBelOBAC11"
/note="end : SP6"

ORIGIN
Query Match 16.7%; Score 49; DB 10; Length 987;
Best Local Similarity 32.4%; Pred. No. 1.4;
Matches 72; Conservative 55; Mismatches 95; Indels 0; Gaps 0;

QY      64 AATTATTTCTCCAGTGGTTTGTGTTATCTCTCAATGTTATAGTGAATAATAATAAT 123
Db      913 WWWWWHTTTHWWTTTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWT 854

QY      124 GCAGAAAAAGCATAAATCAACTATTGAATAACGAAAAATGTGAACCTCTCTCTCTCT 183
Db      853 TTTTATAWAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 794

QY      184 ACTCAGGCTGCCTACCGCTTTTAAATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 243
Db      793 TAATTAATAWAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 734

QY      244 CTAAAGATGTTTCAGAGTAAGATAAATAACAAAGAGAGTAAT 285
Db      733 TTTWTTTTWTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 692

RESULT 11
CNS015FK
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence SP6 end of BAC
BACN13G14 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION
VERSION AL105386
AL105386.1 GI:5617400
KEYWORDS
SOURCE
ORGANISM
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 897)
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billaud at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelOBAC11.

```

collaboration with the European Drosophila Genome Project (EDGP) - <http://www.edgp.ebi.ac.uk> - This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelobAC11.

# FEATURES

Location/Qualifiers  
 1. .897  
 /organism="Drosophila melanogaster"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:7227"  
 /clones="BACN13G14"  
 /clone\_lib="DrosBAC"  
 /plasmid="pBelobAC11"  
 /note="end : SP6"

# ORIGIN

Query Match 16.4%; Score 48.2; DB 10; Length 897;  
 Best Local Similarity 32.2%; Pred. No. 2.1;  
 Matches 93; Conservative 75; Mismatches 119; Indels 2; Gaps 1;  
 Qy 8 AAGGTCCTCGACAGTGTCTCAAGAGATCAGGATTTGAGATCTCTTACTTCAAAATT 67  
 Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 657  
 598 DAKBKCKBTGDATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 657  
 Qy 68 ATTCTCCAGGCTTTCTTCTCTCAATGTTATAGTGA--AATATAATAATACATGC 125  
 Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 717  
 658 MWTMHTYAKTWYTCATBNWYATAMAATTWYATATMAAYTTTATAAATAAABMATTC 717  
 Qy 126 AGAAAAAGCAATAATCAACTATTGAATAACGAAATGTGAACCTCTATGTAACTGCAAC 185  
 Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 777  
 718 ATAHATAAATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 777  
 Qy 186 TCAGGCTGCCCTACCGTTTAAATTTCTCTCTAACTCTTAAATAAATAAATAAATAAATAA 245  
 Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 837  
 778 HCAAGCKCTCTTCAMTTTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 837  
 Qy 246 AAAGATGTTTCAGAGTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 294  
 Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 886  
 838 MTMTAYACTCAWYAYWMMHAATAYATWATTYTAATAAATAAATAAATAAATAAATAA 886

RESULT 12  
 CNS0161W/c 979 bp DNA linear GSS 26-JUL-1999  
 LOCUS Drosophila melanogaster genome survey sequence SP6 end of BAC  
 DEFINITION BACN15917 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.  
 ACCESSION AL106190.1 GI:5620578  
 VERSION AL106190.1  
 KEYWORDS GSS.  
 SOURCE Drosophila melanogaster (fruit fly)  
 ORGANISM Drosophila melanogaster  
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
 REFERENCE 1 (bases 1 to 979)  
 AUTHORS Genoscope.  
 TITLE Direct Submission  
 JOURNAL Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : [seqref@genoscope.cns.fr](mailto:seqref@genoscope.cns.fr))  
 COMMENT - Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)  
 Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - <http://www.edgp.ebi.ac.uk> - This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelobAC11.

# FEATURES

Location/Qualifiers  
 1. .979

/organism="Drosophila melanogaster"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:7227"  
 /clones="BACN15B17"  
 /clone\_lib="DrosBAC"  
 /plasmid="pBelobAC11"  
 /note="end : SP6"

# ORIGIN

Query Match 16.3%; Score 48; DB 10; Length 979;  
 Best Local Similarity 35.2%; Pred. No. 2.3;  
 Matches 81; Conservative 48; Mismatches 101; Indels 0; Gaps 0;  
 Qy 56 TTACTTCAAAATTATCTTCAGTGGTTTCTTCTCAATGTTATAGTGAATATAA 115  
 Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 919  
 978 TTTTATAAATWTTTTTTHHTTATTTCTCTTTATTTTATTTTATTTTATTTTATTTTATTT 919  
 Qy 116 TAATACATGCAGAAAAAGCATAAATCAACTATTGAATAACGAAAAATGTGAACCTCTATG 175  
 Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 859  
 918 AATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 859  
 Qy 176 TAATGCAACTCAGGCTGCCTACCGTTTAAATTTCTCTCTCTCTCTCTCTCTCTCTCTCT 235  
 Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 799  
 858 TTNTNMTATYGTWTTTTTWTWATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 799  
 Qy 236 TATCTTTACTAAAGATGTTTTCAGAGTAAATAAATAAATAAATAAATAAATAAATAAATAA 285  
 Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 749  
 798 TMTTDTCT 749

RESULT 13  
 CNS016LI/c 1101 bp DNA linear GSS 26-JUL-1999  
 LOCUS Drosophila melanogaster genome survey sequence T7 end of BAC  
 DEFINITION BACN16D22 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.  
 ACCESSION AL106896  
 VERSION AL106896.1 GI:5624374  
 KEYWORDS GSS.  
 SOURCE Drosophila melanogaster (fruit fly)  
 ORGANISM Drosophila melanogaster  
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
 REFERENCE 1 (bases 1 to 1101)  
 AUTHORS Genoscope.  
 TITLE Direct Submission  
 JOURNAL Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : [seqref@genoscope.cns.fr](mailto:seqref@genoscope.cns.fr))  
 COMMENT - Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)  
 Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - <http://www.edgp.ebi.ac.uk> - This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelobAC11.

# FEATURES

Location/Qualifiers  
 1. .1101  
 /organism="Drosophila melanogaster"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:7227"  
 /clones="BACN16D22"  
 /clone\_lib="DrosBAC"  
 /plasmid="pBelobAC11"  
 /note="end : T7"

Query Match 16.3%; Score 48; DB 10; Length 1101;  
 Best Local Similarity 36.5%; Pred. No. 2.3;  
 Matches 92; Conservative 55; Mismatches 104; Indels 1; Gaps 1;



```

Qy 43 ATTGAGACTCTCTTACTTCACAAATATTCTCCAGTGGTTTGTGTATCTTCTCAATGTTA 102
   ||: : ||: : ||: : ||: : ||: : ||: : ||: : ||: : ||: : ||: : ||: :
Db 1019 ANTWYWCATTAHWWAWWTTTATATWAATHATTWATAHTATYTWTHWATYTWYTWY 960
   ||: : ||: : ||: : ||: : ||: : ||: : ||: : ||: : ||: : ||: : ||: :
Qy 103 TAGGAAATATAATAATACATGCAGAAAAAGCATAAATCAACTATTGTAATACGAAAAAT 162
   ||: : ||: : ||: : ||: : ||: : ||: : ||: : ||: : ||: : ||: : ||: :
Db 959 TATWAWAHAHAATAATAAAYWHATAATWAWATAAATTTWMT-WTAWMTWTWMTAAT 901
   ||: : ||: : ||: : ||: : ||: : ||: : ||: : ||: : ||: : ||: : ||: :
Qy 163 GTGACCTCTATGTAAGTCACTCAGCTGCGCTACCGTCTTTTAAATTTCTCTCAATCTT 222
   ||: : ||: : ||: : ||: : ||: : ||: : ||: : ||: : ||: : ||: : ||: :
Db 900 ATWATWAAATTAATWWWAWHAWMTAYAWWWWATATWMTATATWATAATWAWATTWTAAT 841
   ||: : ||: : ||: : ||: : ||: : ||: : ||: : ||: : ||: : ||: : ||: :
Qy 223 TAAAAACAATACTATCTTTACTAAAGATGTTTCAGAGTAAGATAAATAACAAGAGAGT 282
   ||: : ||: : ||: : ||: : ||: : ||: : ||: : ||: : ||: : ||: : ||: :
Db 840 TATATWAAATATATATWTAWAAAAAAAWTHAAWAAWAAATATWAAAAAAAHMTAAAA 781
   ||: : ||: : ||: : ||: : ||: : ||: : ||: : ||: : ||: : ||: : ||: :
Qy 283 AATTCACAAAAA 294
   ||: : ||: : ||: : ||: : ||: : ||: : ||: : ||: : ||: : ||: : ||: :
Db 780 WATAAAAAVAAA 769
   ||: : ||: : ||: : ||: : ||: : ||: : ||: : ||: : ||: : ||: : ||: :

RESULT 14
CNS01219 1101 bp DNA linear GSS 26-JUL-1999
LOCUS Drosophila melanogaster genome survey sequence SP6 end of BAC
DEFINITION BACN08024 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL101595
VERSION AL101595.1 GI:5613206
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
http://www.edgp.ebi.ac.uk -- This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billaud at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelobAC11.

FEATURES             Location/Qualifiers
     source           1..1101
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /clone="BACN08024"
                     /clone_lib="DrosBAC"
                     /plasmid="pBelobAC11"
                     /note="end : SP6"

ORIGIN
Query Match      16.3%; Score 47.8; DB 10; Length 1101;
Best Local Similarity 41.2%; Pred. No. 2.5;
Matches 82; Conservative 30; Mismatches 87; Indels 0; Gaps 0;

Qy 96 AATGTTATAGTGAATATAATAATACATGCAGAAAAAGCATCAATCAACTATTGTAATAA 155
   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
Db 177 AATTTATATAAAGGAATAATAWAAAGAAATAATWATAWAAAAWAAAAWATAA 236
   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
Qy 156 CGAAAAATGTGAACCTCTATGTAACTGCAACTCAGCTGCGCTACCGTTTAAATTTCTCTC 215
   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
Db 237 AAAAAAAGGAAAAAGTAAWAWMTAAAWAAWTAATAAATTAATWTTWAWAAAAAATTT 296
   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
Qy 216 TAACTTTAAAAACAATACTATCTTTTACTAAAGATGTTTCAGAGTAGTAAGATAAATAACA 275
   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:

```

```

Db 297 WWTWTWTAAWWAAAAAATTTAAAAAAWAAWATAGAAAAATWAAWAAAAAAA 356
   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
Qy 276 AGAGAGTAATTCACAAAAA 294
   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
Db 357 AAAAAAATAAATAAAWAAAMM 375
   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:

RESULT 15
CX572895/c 572 bp mRNA linear EST 13-JAN-2005
LOCUS TTE00022498 Amplicon Express - Conjugative Form Tetrahymena
DEFINITION thermophila cDNA, mRNA sequence.
ACCESSION CX572895
VERSION CX572895.1 GI:57656073
KEYWORDS EST.
SOURCE Tetrahymena thermophila
ORGANISM Tetrahymena thermophila
Eukaryota; Alveolata; Ciliophora; Oligohymenophorea;
Hymenostomatida; Tetrahymenina; Tetrahymenidae; Tetrahymena.
REFERENCE 1 (bases 1 to 572)
AUTHORS Garg, J., Pearlman, R.E. and Carlton, J.
TITLE PEPdbPub (http://amoebidia.bcm.umontreal.ca/public/pepdb/agrm.php)
JOURNAL Tetrahymena thermophila (TIGR)
COMMENT Unpublished (2004)
CONTACT: PEPdb
DEPARTMENT de Biochimie, Universite de Montreal
EMAIL: pepdb-curator@bch.umontreal.ca
PLATE: 3435.

FEATURES             Location/Qualifiers
     source           1..572
                     /organism="Tetrahymena thermophila"
                     /mol_type="mRNA"
                     /db_xref="taxon:5911"
                     /clone_lib="Amplicon Express - Conjugative Form"

ORIGIN
Query Match      16.1%; Score 47.2; DB 8; Length 572;
Best Local Similarity 52.8%; Pred. No. 3.4;
Matches 103; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

Qy 77 GTGGTTTGTGTATCTCTCAATGTTATAGTGAATAATAATAACATGCAGAAAAAGCA 136
   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
Db 572 GTTTTGTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 513
   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
Qy 137 TAAATCAACTATTGAAATAACGAAAAATGTGAACCTCTATGTAACTGCAACTCAGGCTGCCT 196
   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
Db 512 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 453
   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
Qy 197 ACCGTTTTTAATTTCTCTCTAATCTTTAAAAACAATACTATCTTTTACTTAAAGATGTTTC 256
   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
Db 452 ACAATTTCTGTGTTAAATAAATAATATATATTTTATTTTATTTTAAATATTCCTCAATT 393
   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
Qy 257 AGAGTAAGATAAAAAA 272
   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
Db 392 CTTGAACACGCAATA 377
   |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:

Search completed: December 7, 2005, 09:55:45
Job time : 3359.89 secs

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 10:04:41 ; Search time 128.321 Seconds  
(without alignments)  
4072.612 Million cell updates/sec

Title: US-09-980-046B-2

Perfect score: 294

Sequence: 1 cggccttaaggcttcctctga.....aagagagtaattcccaaaaa 294

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.\*  
1: /cgn2\_6/ptodata/1/ina/1\_COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5\_COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/H\_COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/PP\_COMB.seq.\*  
7: /cgn2\_6/ptodata/1/ina/RE\_COMB.seq.\*  
8: /cgn2\_6/ptodata/1/ina/RE\_COMB.seq.\*  
9: /cgn2\_6/ptodata/1/ina/backfile1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	288.2	98.0	84425	3	US-09-949-016-17402
2	43.6	14.8	65415	3	US-09-949-016-16699
3	43	14.6	1664976	3	US-08-916-421B-1
4	43	14.6	1664976	3	US-09-692-570-1
5	42.4	14.4	601	3	US-09-949-016-198281
6	42.4	14.4	767677	3	US-09-949-016-12147
7	42.4	14.4	767677	3	US-09-949-016-17361
8	42	14.3	601	3	US-09-949-016-198280
9	40.8	13.9	613	2	US-08-630-822A-98
10	40.8	13.9	613	2	US-09-005-069-98
11	40.8	13.9	613	3	US-09-171-156A-47
12	40.8	13.9	613	3	US-09-004-730A-47
13	40.8	13.9	613	3	US-08-981-799A-47
14	40.6	13.8	832	3	US-09-621-976-2813
15	40.4	13.7	50000	3	US-09-662-254B-23
16	40.2	13.7	601	3	US-09-949-016-201502
17	40.2	13.7	601	3	US-09-949-016-201726
18	40.2	13.7	2680	3	US-09-949-016-5657
19	40.2	13.7	3349	3	US-09-949-016-17399
20	40.2	13.7	52174	3	US-09-949-002-573
21	40.2	13.7	52289	3	US-09-949-002-720
22	40.2	13.7	86639	3	US-09-949-016-17397
23	39.8	13.5	601	3	US-09-949-016-201501
24	39.8	13.5	601	3	US-09-949-016-201727

c	25	39.4	13.4	32392	3	US-09-662-254B-27	Sequence 27, Appl
c	26	39.4	13.4	50000	3	US-09-662-254B-23	Sequence 23, Appl
c	27	39	13.3	601	3	US-09-949-016-185650	Sequence 185650,
c	28	39	13.3	1052	3	US-09-976-594-819	Sequence 819, App
c	29	39	13.3	168104	3	US-09-949-016-12026	Sequence 12026, A
c	30	39	13.3	168105	3	US-09-949-016-16554	Sequence 16554, A
c	31	38.8	13.2	601	3	US-09-949-016-170445	Sequence 170445, A
c	32	38.4	13.1	97376	3	US-09-949-016-16093	Sequence 16093, A
c	33	38.4	13.1	387902	3	US-09-949-016-14543	Sequence 14543, A
c	34	38.4	13.1	421883	3	US-09-949-016-12557	Sequence 12557, A
c	35	38	12.9	314798	3	US-09-949-016-13539	Sequence 13539, A
c	36	37.8	12.9	16573	3	US-09-949-016-14876	Sequence 14876, A
c	37	37.8	12.9	29133	3	US-09-949-016-12694	Sequence 12694, A
c	38	37.8	12.9	38479	3	US-09-949-016-16730	Sequence 16730, A
c	39	37.6	12.8	3942	3	US-09-601-198-50	Sequence 50, Appl
c	40	37.6	12.8	5448	3	US-09-620-312D-246	Sequence 246, App
c	41	37.6	12.8	5532	3	US-09-620-312D-245	Sequence 245, App
c	42	37.4	12.7	601	3	US-09-949-016-185649	Sequence 185649,
c	43	37.4	12.7	2877	3	US-09-235-103-1	Sequence 1, Appli
c	44	37.4	12.7	113966	3	US-09-949-016-12277	Sequence 12277, A
c	45	37.4	12.7	113967	3	US-09-949-016-17051	Sequence 17051, A

ALIGNMENTS

RESULT 1

US-09-949-016-17402  
; Sequence 17402, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: Fast-Seq for Windows Version 4.0  
; SEQ ID NO 17402  
; LENGTH: 84425  
; TYPE: DNA  
; ORGANISM: Human  
; US-09-949-016-17402

Query Match	98.0%	Score 288.2	DB 3	Length 84425
Best Local Similarity	99.0%	Pred. No. 3.8e-61		
Matches 290	Conservative 0	Mismatches 3	Indels 0	Gaps 0
Qy	1	CGGCCTTAAGGTTCCCTGACAGTTGTCTCAAGAAGATCAGGATTTGAGCATCTCTTACT	60	
Db	53140	CGGCCTTAAGGTTCCCTGACAGTTGTCTCAGAGAATCAGGATTTGAGCATCTCTTACT	53199	
Qy	61	TCAAAATATTCTTCCAGTGGTTTTGTATCTTCTCAATGTTATAGTGAATATAATAATA	120	
Db	53200	TCAAAATATTCTTCCAGTGGTTTTGTATCTTCTCAATGTTATAGTGAATATAATAATA	53259	
Qy	121	CATGCGAAAAAGCATATAATCAACTATTGAATACGAAATGGAACCTCTATGTAAT	180	
Db	53260	CATGCGAAAAAGCATATAATCAACTATTGAATACGAAATGGAACCTCTATGTAAT	53319	
Qy	181	GCAACTCAGGCTGCCTACCGTTTTTAATTTCTCTCTAATCTTTAAAAACAATACTATCT	240	
Db	53320	GCAACTCAGGCTGCCTACCGTTTTTAATTTCTCTCTAATCTTTAAAAACAATACTATCT	53379	
Qy	241	TTACTAAAGATGTTTTCAGAGTAAGATAAAAAATACAAAGAGAGTAATTCCTCCAAA	293	

```
Db 53380 TTACTAAGATGTTTCAGAGTAAGATAAAATACAAAGAGAGTAATTCCTCAA 53432

RESULT 2
US-09-949-016-16699/c
; Sequence 16699, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16699
; LENGTH: 65415
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16699

Query Match 14.8%; Score 43.6; DB 3; Length 65415;
Best Local Similarity 48.8%; Pred. No. 0.27;
Matches 118; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

QY 2 GGCCTTAAGTTCCTCGACAGTGTCTCAAGAGATCAGAGATTGAGCATCTTACTTT 61
Db 39198 GGCATCAAGTCCTCTCGATGGTTTCTTTTCAGTTCCTTACTTTATTTAGATCTTCT 39139
QY 62 CAAATATTCTTCAGTGGTGGTTTGTATCTCTCAATGTTATAGTGAATATAATATAC 121
Db 39138 CAGACTCGAATTAAGTCTGGTGTAGTCAAGATGCAATTTGATCTTTAAAAATGGA 39079
QY 122 ATGCAGAAAAAGCATAAATCAACTATTGTAATACGAAAAATGGAACCTCTATGTAAC 181
Db 39078 GGAAGAAGAAGATTAAATATATAGACACATTTTTCAAAGTTTGAAGTCCAATGTA 39019
QY 182 CAACTAGGCTGCTACCGTTTAAATTTCTCTAATCTTTAAACAAATATATCTTT 241
Db 39018 TATGTTACTGCTTTAATAACTTTTTTTTGAATACTCTAAAAATACCCATATAC 38959
QY 242 TA 243
Db 38958 CA 38957

RESULT 3
US-08-916-421B-1
; Sequence 1, Application US/08916421B
; Patent No. 6503729
; GENERAL INFORMATION:
; APPLICANT: Bult et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
; jannaschii
; Patent No. 6503729
; FILE REFERENCE: PR275
; CURRENT APPLICATION NUMBER: US/08/916,421B
; CURRENT FILING DATE: 1997-08-22
; PRIOR FILING DATE: 1997-08-22
; PRIOR FILING DATE: 1996-08-22
; PRIOR FILING DATE: 1996-08-22
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
```

```
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28222)..(28222)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (28257)..(28258)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84773)..(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84808)..(84808)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84812)..(84812)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98120)..(98120)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98159)..(98159)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98239)..(98239)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98266)..(98266)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98343)..(98343)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (103998)..(103998)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (148948)..(148948)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (163385)..(163385)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (191989)..(191989)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (191995)..(191995)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (231980)..(231980)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (234187)..(234187)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (234220)..(234220)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (234814)..(234814)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (309398)..(309398)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (309418)..(309418)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (312837)..(312837)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (312993)..(312993)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (319226)..(319226)
; OTHER INFORMATION: n equals a, t, c, or g
```

```

; LOCATTION: (1602912)..(1602912)
; NAME/KEY: n equals a, t, c, or g
; OTHER INFORMATION: n equals a, t, c, or g
; LOCATTION: (1603734)..(1603734)
; NAME/KEY: n equals a, t, c, or g
; OTHER INFORMATION: n equals a, t, c, or g
; LOCATTION: (1637998)..(1637998)
; NAME/KEY: n equals a, t, c, or g
; OTHER INFORMATION: n equals a, t, c, or g
; LOCATTION: (1664855)..(1664855)
; NAME/KEY: n equals a, t, c, or g
; OTHER INFORMATION: n equals a, t, c, or g
US-08-916-421B-1

Query Match      14.6%; Score 43; DB 3; Length 1664976;
Best Local Similarity 52.5%; Pred. No. 0.68;
Matches 94; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

Qy      84  TGTATCTCTCTCAATGTTATAGTGAATATAATAATCATGACAGAAAAAGCATAAATCA 143
Db      411276  TGTCTTTTCATAAAGTAGGATAAATTCATTAAATCTCCATCTATATATCTCAATAA 411335

Qy      144  ACTATTGAATAAACGAAAATGTGAACCTCTATGTAACTGCAACTCAGGCTGCCTACCGTTT 203
Db      411336  ACTCTATTAAAAATAGCTATAGCCCTCATTAACCTAATCTATAGCTCTTAAACGGATT 411395

Qy      204  TTAATTTCTCTCTAATCTTTAAAAACAAATACTATCTTTACTAAAGATGTTTTCAGAGTA 262
Db      411396  TTTTACCCTTACTTTAACTTTAAATTTTATACCTCTCTTTCTTTAAATGTTTATAAACTA 411454

RESULT 4
US-09-692-570-1
; Sequence 1, Application US/09692570
; Patent No. 6797466
; GENERAL INFORMATION:
; APPLICANT: Bult et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Meth
; Patent No. 6797466
; TITLE OF INVENTION: jannaschii
; FILE REFERENCE: PB275C1
; CURRENT APPLICATION NUMBER: US/09/692,570
; CURRENT FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: US 60/024,428
; PRIOR FILING DATE: 1996-08-22
; PRIOR APPLICATION NUMBER: US 08/916,421
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (28222)..(28222)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (28257)..(28258)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (84773)..(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (84808)..(84808)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (84812)..(84812)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:

```

```
;/ NAME/KEY: misc feature
;/ LOCATION: (98120)..(98120)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (98159)..(98159)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (98239)..(98239)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (98266)..(98266)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (98343)..(98343)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (103998)..(103998)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (148948)..(148948)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (163385)..(163385)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (191989)..(191989)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (191995)..(191995)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (231980)..(231980)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (234187)..(234187)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (234220)..(234220)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (234814)..(234814)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (309398)..(309398)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (309418)..(309418)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (312837)..(312837)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (312993)..(312993)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
```

```
;/ LOCATION: (319226)..(319226)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (559167)..(559167)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (559241)..(559241)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (600992)..(600992)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (622708)..(622708)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (657081)..(657081)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (657203)..(657203)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (674435)..(674435)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (682442)..(682442)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (713652)..(713652)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (741684)..(741684)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (779455)..(779455)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (779676)..(779676)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (855539)..(855539)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (871619)..(871619)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (1084830)..(1084830)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (1096846)..(1096846)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (1119881)..(1119881)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (1130881)..(1130881)
```

OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1310988)..(1310988)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1313224)..(1313224)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1349473)..(1349473)  
OTHER INFORMATION: n equals a, t, c, or g

Query Match 14.4%; Score 43; DB 3; Length 1664976;  
Best Local Similarity 52.5%; Pred. No. 0.68;  
Matches 94; Conservative 0; Mismatches 85; Indels 0; Gaps 0;  
Qy 84 TGTTCCTCTCTCAATGTTATAGTGAATATAATATAATACATGCAGAAAAAGCATATAATCA 143  
Db 411276 TGTCTTTTCATAAAGTAGGATAAATTACATTAATAATCTCCATCTATATATCTCAATAA 411335  
Qy 144 ACTATTGAAATACGAAATGTGAACCTCTATGTAACTGCAACTCAGGCTGCTACCGTTT 203  
Db 411336 ACTCTCATTTAAATAAAGCTATAGCCCTCATTAACCTAACTCTATAGCTCTAAACGGATT 411395  
Qy 204 TTAATTTCTCTCTAATCTTTTAAACAAATACTATCTTTACTAAAGATGTTTCAGAGTA 262  
Db 411396 TTTTACCTTACTTAACTTTAAATTTTATACCTCTCTTTTAAATAGTTTATAAATA 411454

RESULT 5  
US-09-949-016-198281/c  
; Sequence 198281, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 198281  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-198281

Query Match 14.4%; Score 42.4; DB 3; Length 601;  
Best Local Similarity 50.0%; Pred. No. 0.23;  
Matches 106; Conservative 0; Mismatches 106; Indels 0; Gaps 0;  
Qy 44 TTTGAGCATCTCTTACTTCAAATATTCTCCAGTGGTTTTGTTATCTTCTCAATGTTAT 103  
Db 315 TTTAAGTAACCTCATGCTTAAACTATATTTTCATATGTCAAAGGGATATGTTTGAAGCTAG 256  
Qy 104 AGTGAATATAATAATACATGCAGAAAAAGCATATACTATGTAATGAATACGAAATG 163  
Db 255 AGTCAAAATCTTTTCTCATGTAAGAATGTGTAATGATTATATTAATAAAAAAATG 196  
Qy 164 TGAACCTCTATGTAACCTGCAACTCAGGCTGCTACCGTTTTTAATTTCTCTCAATCTTT 223  
Db 195 TTGTAGAACAAATATTTGCAAAATACATTTGTTTCATATTTGTTCAATAAGTTATATAAAGT 136  
Qy 224 AAAAACAAATACTATCTTTTACTAAAGATGTTT 255

Db 135 AACAGAAATATAGTTTTTTTCCAACTTCCTTT 104

## RESULT 6

US-09-949-016-12147  
; Sequence 12147, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12147  
; LENGTH: 767677  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)..(767677)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-12147

Query Match 14.4%; Score 42.4; DB 3; Length 767677;  
Best Local Similarity 50.0%; Pred. No. 0.83; 106; Indels 0; Gaps 0;  
Matches 106; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

Qy 44 TTTGAGCATCTCTTACTTCAAATATTCTCCAGTGGTTTTGTTATCTTCTCAATGTTAT 103  
Db 668111 TTTAAGTAACCTCATGCTTAAACTATATTTTCATATGTCAAAGGGATATGTTTGAAGCTAG 668170  
Qy 104 AGTGAATATAATAATACATGCAGAAAAAGCATATACTATGTAATGAATACGAAATG 163  
Db 668171 AGTCAAAATCTTTTCTCATGTAAGAATGTGTAATGATTATATTAATAAAAAAATG 668230  
Qy 164 TGAACCTCTATGTAACCTGCAACTCAGGCTGCTACCGTTTTTAATTTCTCTCAATCTTT 223  
Db 668231 TTGTAGAACAAATATTTGCAAAATACATTTGTTTCATATTTGTTCAATAAGTTATATAAAGT 668290  
Qy 224 AAAAACAAATACTATCTTTTACTAAAGATGTTT 255  
Db 668291 AACAGAAATATAGTTTTTTTCCAACTTCCTTT 668322

## RESULT 7

US-09-949-016-17361  
; Sequence 17361, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0

```

; SEQ ID NO 17361
; LENGTH: 767677
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_
; LOCATION: (1)..
; OTHER INFORMATION:
US-09-949-016-17361

```

Query Match	14.4%	Score 42.4	DB 3	Length 767677
Best Local Similarity	50.0%	Pred. No. 0.83		
Matches 106	Conservative 0	Mismatches 106	Indels 0	Gaps 0
Qy	44	TTTGAGCATCTCTTACTTCAAAATATTCTTCAGTGGTTTTGGTTATCTTCTCAATGTTAT	103	
Db	668111	TTTAAAGTAACTCATTTGCTTAAACTATATTTCATATGTCAAAGGATATGTTTCGAGCTAG	668170	
Qy	104	AGTGAATATATAATATCATCGACAGAAAAAGCATAAATCACTATTGTAATACGAAAATG	163	
Db	668171	AGTCAAAATCTTTTTCTCATTTGAATAAAGAAATGTGTAITGATTTATTAATAAAAAAATG	668230	
Qy	164	TGAACCTCTATGTAACGTGCAACTCAGGCTGCCTACCGTTTTTAAITTTCTCTCTAATCTTT	223	
Db	668231	TTGTAGAACAAATATTTGCAAAATACATTTGTTTCATATTGTTTCAAAATAGTTATAAAAAGT	668290	
Qy	224	AAAAACAAATACTACTTTTACTAAAGATGTTTT	255	
Db	668291	AAAACAGAAATATAGTTTTTTTCCAACTTCTCTTT	668322	

```

RESULT 8
US-09-949-016-198280/c
; Sequence 198280, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 198280
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-198280

```

Query Match	14.3%	Score 42	DB 3	Length 601
Best Local Similarity	49.5%	Pred. No. 0.29		
Matches 105	Conservative 1	Mismatches 106	Indels 0	Gaps 0
QY	44	TTTGAGCATCTCTTACTTCAAAATTATCTTCCAGTGGTTTTGGTTATCTTCTCAATGTTAT	103	
DB	387	TTTAAGTAACTCATGTGCTTAACTATATTTTCATATGTCAAAGGATATGTTCTGAAGCTAG	328	
QY	104	AGTGAATAATAATAATCATGCAGAAAAAGCATAAATCAACTATTGTAATACGAAAAATG	163	
DB	327	AGTCAAAATCTTTTCTCATTGAAATAAGAAATGTGTATTGATTTATATTAATAAAAAAATG	268	
QY	164	TGAACCTCTATGTAACTGCAACTCAGGCTGCCTACCGTTTTTAATTTCTCTCTAATCTTT	223	
DB	267	TTGTAGAACCAAAATATTTGCAAAATACATTTGTTTCATATTTGTTCAAAATGAATTTATAAAAAGT	208	
QY	224	AAAAACAAATACTATCTTTTACTAAAGATGTTT	255	

Db 207 AACAGAAATATAGTTTTTCCAACTTCCTTT 176

RESULT 9  
 US-08-630-822A-98/C  
 Sequence 98, Application US/08630822A  
 Patent No. 5840695  
 GENERAL INFORMATION:  
 APPLICANT: FRANK, GLENN R.  
 APPLICANT: HUNTER, SHIRLEY WU  
 APPLICANT: WALLENFELS, LYNDA  
 TITLE OF INVENTION: NOVEL ECOTOPARASITE SALIVA PROTEINS  
 TITLE OF INVENTION: AND APPARATUS TO COLLECT SUCH PROTEINS  
 NUMBER OF SEQUENCES: 107  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Sheridan Ross P.C.  
 STREET: 1700 Lincoln Street, Suite 3500  
 CITY: Denver  
 STATE: Colorado  
 COUNTRY: U.S.A.  
 ZIP: 80203  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/630,822A  
 FILING DATE: 11-APR-1996  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: CONNELL, GARY J.  
 REGISTRATION NUMBER: 32,020  
 REFERENCE/DOCKET NUMBER: 2618-17-C3  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (303) 863-9700  
 TELEFAX: (303) 863-0223  
 INFORMATION FOR SEQ ID NO: 98:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 613 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 US-08-630-822A-98

```

Query Match      13.9%;      Score 40.8;  DB 2;  Length 613;
Best Local Similarity 52.3%;  Pred. No. 0.59;
Matches 90;  Conservative 0;  Mismatches 82;  Indels 0;  Gaps 0;

Qy      67  TATCTTCGAGTGGTGTGTTTATCTTCTCAATGTTATAGTGAATATAATAATACATGCA 126
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      613  TTTTITTTTTTTTTTTTTTTTTTAAATCTTATATAATATTTATTATCATATAAATACAAAGTTCC 554

Qy      127  GAAAAAGCATAAATCAACTATTGCAATAACGAAATGTGAACCTCTATGTAACTGCACT 186
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      553  AAAAGTATTAAATATAAAATTCATCAATTTATTGTAAGATAAACATTTTTTAAGTAAATATCAAG 494

Qy      187  CAGGCTGCCTACCGTTTTTAAATTTCTCTCTAAATCTTTTAAAAACAAATACTAT 238
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      493  TCATCTCAACATCGTTTTTCAATTTCATCTTATATTTTAAATTTAAGTAATAT 442

RESULT 10
US-09-005-069-98/c
; Sequence 98, Application US/09005069
; Patent No. 5932470
; GENERAL INFORMATION:
; APPLICANT: FRANK, GLENN R.
; APPLICANT: HUNTER, SHIRLEY WU
; APPLICANT: WALLENFELS, LYNDA
; TITLE OF INVENTION: NOVEL ECTOPARASITE SALIVA PROTEINS

```







GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 10:04:41 ; Search time 34.4809 Seconds  
(without alignments)  
4072.612 Million cell updates/sec

Title: US-09-980-046b-3  
Perfect score: 79  
Sequence: 1 cgggaatcccccttcagttc.....aaacaacgactctaaaaa 79

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA.\*  
1: /cgn2\_6/ptodata/1/ina/1 COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5 COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/H COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/PPUS COMB.seq.\*  
7: /cgn2\_6/ptodata/1/ina/PP COMB.seq.\*  
8: /cgn2\_6/ptodata/1/ina/RE COMB.seq.\*  
9: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	74.8	94.7	601	3	US-09-949-016-26334, A
C 2	74.8	94.7	601	3	US-09-949-016-105203, A
C 3	74.8	94.7	2000	3	US-09-997-165-3, Appli
4	74.8	94.7	2000	3	US-09-949-016-426, App
5	74.8	94.7	2001	3	US-09-949-016-2936, App
6	74.8	94.7	2180	2	US-08-755-559-2, Appli
7	74.8	94.7	2180	3	US-09-210-474-2, Appli
8	74.8	94.7	2180	3	US-09-539-774-2, Appli
9	74.8	94.7	16738	3	US-09-949-016-12168, A
10	74.8	94.7	16738	3	US-09-949-016-14678, A
11	28.2	35.7	45983	3	US-09-949-016-17010, A
12	27.8	35.2	818128	3	US-09-949-016-14546, A
13	27.8	35.2	818128	3	US-09-949-016-14547, A
14	27.8	35.2	818128	3	US-09-949-016-14548, A
15	27.8	35.2	818128	3	US-09-949-016-14549, A
16	27.8	35.2	818128	3	US-09-949-016-14550, A
17	27.8	35.2	818128	3	US-09-949-016-14551, A
18	27.8	35.2	818128	3	US-09-949-016-14552, A
19	27.8	35.2	818128	3	US-09-949-016-14553, A
20	27.8	35.2	818128	3	US-09-949-016-14554, A
21	27.8	35.2	818128	3	US-09-949-016-14555, A
22	27.8	35.2	818128	3	US-09-949-016-14556, A
23	27.8	35.2	818128	3	US-09-949-016-14557, A
24	27.8	35.2	818128	3	US-09-949-016-14558, A

ALIGNMENTS

RESULT 1

US-09-949-016-26334/c  
; Sequence 26334, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 26334  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-26334

Query Match 94.7%; Score 74.8; DB 3; Length 601;  
Best Local Similarity 97.4%; Pred. No. 3e-15; 2; Indels 0; Gaps 0;  
Matches 76; Conservative 0; Mismatches 2;

Qy 1 CGGGAATCCCCCTTCAGTCTTTTGAAGAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60  
Db 129 CGGGAATCCCCCTTCAGTCTTTTGAAGAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 70  
Qy 61 AACAAACCGACTCTAAAA 78  
Db 69 AACAAACCGACTCAGAAA 52

RESULT 2

US-09-949-016-105203/c  
; Sequence 105203, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307

```
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 105203
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-105203

Query Match          94.7%; Score 74.8; DB 3; Length 601;
Best Local Similarity 97.4%; Pred. No. 3e-15;
Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCGCTTCAGTCTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
Db 129 CGGGAATCCCGCTTCAGTCTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 70
Qy 61 AACAAACCGACTCTAAAA 78
Db 69 AACAAACCGACTCAAAA 52

RESULT 3
US-09-997-165-3
; Sequence 3, Application US/09997165
; Patent No. 6762030
; GENERAL INFORMATION:
; APPLICANT: Lyman, Stewart D.
; APPLICANT: Fanslow, William C.
; TITLE OF INVENTION: LIGAND FOR CD7 AND METHODS OF USE THEREOF
; FILE REFERENCE: 2913-US
; CURRENT APPLICATION NUMBER: US/09/997,165
; CURRENT FILING DATE: 2001-11-27
; PRIOR APPLICATION NUMBER: PCT/US00/14612
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/136,450
; PRIOR FILING DATE: 1999-05-28
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (119)..(865)
US-09-997-165-3

Query Match          94.7%; Score 74.8; DB 3; Length 2000;
Best Local Similarity 97.4%; Pred. No. 4e-15;
Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCGCTTCAGTCTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
Db 1645 CGGGAATCCCGCTTCAGTCTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1704
Qy 61 AACAAACCGACTCTAAAA 78
Db 1705 AACAAACCGACTCAAAA 1722

RESULT 4
US-09-949-016-426
; Sequence 426, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
```

```
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 426
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-426

Query Match          94.7%; Score 74.8; DB 3; Length 2000;
Best Local Similarity 97.4%; Pred. No. 4e-15;
Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCGCTTCAGTCTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
Db 1645 CGGGAATCCCGCTTCAGTCTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1704
Qy 61 AACAAACCGACTCTAAAA 78
Db 1705 AACAAACCGACTCAAAA 1722

RESULT 5
US-09-949-016-2936
; Sequence 2936, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2936
; LENGTH: 2001
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-2936

Query Match          94.7%; Score 74.8; DB 3; Length 2001;
Best Local Similarity 97.4%; Pred. No. 4e-15;
Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCGCTTCAGTCTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
Db 1646 CGGGAATCCCGCTTCAGTCTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1705
Qy 61 AACAAACCGACTCTAAAA 78
Db 1706 AACAAACCGACTCAAAA 1723

RESULT 6
US-08-755-559-2
```

; Sequence 2, Application US/0875559  
; Patent No. 5912142  
; GENERAL INFORMATION:  
; APPLICANT: KAUFMAN, RUSSEL E.  
; APPLICANT: SLENTZ-KESLER, KIMBERLY  
; TITLE OF INVENTION: GENE PRODUCT OVER EXPRESSED IN CANCER  
; TITLE OF INVENTION: CELLS  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHVE P.C.  
; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR  
; CITY: ARLINGTON  
; STATE: VIRGINIA  
; COUNTRY: U.S.A.  
; ZIP: 22201-4714  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/755,559  
; FILING DATE: 22-NOV-1996  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: WILSON, MARY J.  
; REGISTRATION NUMBER: 32,955  
; REFERENCE/DOCKET NUMBER: 1579-116  
; TELEPHONE: (703) 816-4000  
; TELEFAX: (703) 816-4100  
; INFORMATION FOR SEQ ID NO: 2:  
; TYPE: nucleic acid  
; LENGTH: 2180 base pairs  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; US-08-755-559-2

Query Match 94.7%; Score 74.8; DB 2; Length 2180;  
Best Local Similarity 97.4%; Pred. No. 4.1e-15;  
Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60  
Db 1841 CGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1900

Qy 61 AACAAACCGACTCTAAAA 78  
Db 1901 AACAAACCGACTCTACAA 1918

RESULT 7  
US-09-210-474-2  
; Sequence 2, Application US/09210474  
; Patent No. 6072034  
; GENERAL INFORMATION:  
; APPLICANT: KAUFMAN, RUSSEL E.  
; APPLICANT: SLENTZ-KESLER, KIMBERLY  
; TITLE OF INVENTION: GENE PRODUCT OVER EXPRESSED IN CANCER  
; TITLE OF INVENTION: CELLS  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHVE P.C.  
; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR  
; CITY: ARLINGTON  
; STATE: VIRGINIA  
; COUNTRY: U.S.A.  
; ZIP: 22201-4714  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/210,474  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/755,559  
; FILING DATE: 22-NOV-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: WILSON, MARY J.  
; REGISTRATION NUMBER: 32,955  
; REFERENCE/DOCKET NUMBER: 1579-116  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 816-4000  
; TELEFAX: (703) 816-4100  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2180 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; US-09-210-474-2

Query Match 94.7%; Score 74.8; DB 3; Length 2180;  
Best Local Similarity 97.4%; Pred. No. 4.1e-15;  
Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60  
Db 1841 CGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1900

Qy 61 AACAAACCGACTCTAAAA 78  
Db 1901 AACAAACCGACTCTACAA 1918

RESULT 8  
US-09-539-774-2  
; Sequence 2, Application US/09539774  
; Patent No. 6350615  
; GENERAL INFORMATION:  
; APPLICANT: KAUFMAN, RUSSEL E.  
; APPLICANT: SLENTZ-KESLER, KIMBERLY  
; TITLE OF INVENTION: GENE PRODUCT OVER EXPRESSED IN CANCER  
; TITLE OF INVENTION: CELLS  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHVE P.C.  
; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR  
; CITY: ARLINGTON  
; STATE: VIRGINIA  
; COUNTRY: U.S.A.  
; ZIP: 22201-4714  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/539,774  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/210,474  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: WILSON, MARY J.  
; REGISTRATION NUMBER: 32,955  
; REFERENCE/DOCKET NUMBER: 1579-116  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 816-4000

```
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 2180 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: DNA (genomic)
US-09-539-774-2

Query Match          94.7%; Score 74.8; DB 3; Length 2180;
Best Local Similarity 97.4%; Pred. No. 4.1e-15;
Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGGGAATCCCCCTTCAGTTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
Db 1841 CGGGAATCCCCCTTCAGTTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1900

QY 61 AACAAACCGACTCTAAAA 78
Db 1901 AACAAACCGACTCAGAAA 1918

RESULT 9
US-09-949-016-12168
; Sequence 12168, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12168
; LENGTH: 16738
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12168

Query Match          94.7%; Score 74.8; DB 3; Length 16738;
Best Local Similarity 97.4%; Pred. No. 6.7e-15;
Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGGGAATCCCCCTTCAGTTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
Db 14383 CGGGAATCCCCCTTCAGTTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 14442

QY 61 AACAAACCGACTCTAAAA 78
Db 14443 AACAAACCGACTCAGAAA 14460

RESULT 10
US-09-949-016-14678
; Sequence 14678, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
```

```
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14678
; LENGTH: 16738
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14678

Query Match          94.7%; Score 74.8; DB 3; Length 16738;
Best Local Similarity 97.4%; Pred. No. 6.7e-15;
Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGGGAATCCCCCTTCAGTTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
Db 14383 CGGGAATCCCCCTTCAGTTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 14442

QY 61 AACAAACCGACTCTAAAA 78
Db 14443 AACAAACCGACTCAGAAA 14460

RESULT 11
US-09-949-016-17010
; Sequence 17010, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17010
; LENGTH: 45983
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(45983)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17010

Query Match          35.7%; Score 28.2; DB 3; Length 45983;
Best Local Similarity 64.6%; Pred. No. 18;
Matches 42; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 13 TTCAGTTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAAACAAACCGACT 72
Db 12689 TTTTGTCTTTGAAATAGTTCCATGATTTAGCTCTGAAAGGCAAAATATATATCACT 12748

QY 73 CTAAG 77
Db 12749 CTCAG 12753

RESULT 12
US-09-949-016-14546
; Sequence 14546, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
```

APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 14546  
; LENGTH: 818128  
; TYPE: DNA  
; ORGANISM: Human  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(818128)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-14546

Query Match 35.2%; Score 27.8; DB 3; Length 818128;  
Best Local Similarity 65.1%; Pred. No. 50;  
Matches 41; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy 12 CTTGAGTTCCTTGAAGAAAGTCCGACTCGAATATCTGAAATGAAGAAACAAACCGAC 71  
Db 61735 CTTGAGTTCCTGAGCACATCCGAGTTTCTTATCTCTGGAGTGAGACTTCAAAAATAC 61794

Qy 72 TCT 74  
Db 61795 TCT 61797

RESULT 13  
US-09-949-016-14547  
; Sequence 14547, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 14547  
; LENGTH: 818128  
; TYPE: DNA  
; ORGANISM: Human  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(818128)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-14547

Query Match 35.2%; Score 27.8; DB 3; Length 818128;  
Best Local Similarity 65.1%; Pred. No. 50;  
Matches 41; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy 12 CTTGAGTTCCTTGAAGAAAGTCCGACTCGAATATCTGAAATGAAGAAACAAACCGAC 71  
Db 61735 CTTGAGTTCCTGAGCACATCCGAGTTTCTTATCTCTGGAGTGAGACTTCAAAAATAC 61794

Qy 72 TCT 74  
Db 61795 TCT 61797

## RESULT 14

US-09-949-016-14548  
; Sequence 14548, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 14548  
; LENGTH: 818128  
; TYPE: DNA  
; ORGANISM: Human  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(818128)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-14548

Query Match 35.2%; Score 27.8; DB 3; Length 818128;  
Best Local Similarity 65.1%; Pred. No. 50;  
Matches 41; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy 12 CTTGAGTTCCTTGAAGAAAGTCCGACTCGAATATCTGAAATGAAGAAACAAACCGAC 71  
Db 61735 CTTGAGTTCCTGAGCACATCCGAGTTTCTTATCTCTGGAGTGAGACTTCAAAAATAC 61794

Qy 72 TCT 74  
Db 61795 TCT 61797

## RESULT 15

US-09-949-016-14549  
; Sequence 14549, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 14549  
; LENGTH: 818128  
; TYPE: DNA  
; ORGANISM: Human  
; NAME/KEY: misc\_feature  
US-09-949-016-14549

```
! LOCATION: (1)...(818128)
! OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14549

Query Match      35.2%; Score 27.8; DB 3; Length 818128;
Best Local Similarity 65.1%; Pred. No. 50;
Matches 41; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy      12  CTTTCAGTTCCTTGAAGAGTTCCATGACTCGAATATCTGAAATGAGAGAAACAAACCGAC 71
          |||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      61735  CTTTCAGTTCCTTGAAGAGTTCCATGACTCGAATATCTGAAATGAGAGAAACAAACCGAC 71
          |||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

Qy      72  TCT 74
          |||
Db      61795  TCT 61797
```

Search completed: December 6, 2005, 11:47:18  
Job time : 44.4809 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 21:43:57 ; Search time 901.725 Seconds  
(without alignments)  
4099.008 Million cell updates/sec

Title: US-09-980-046B-3

Perfect score: 79

Sequence: 1 cgggaatcccccttcagttc.....aaacaaccgactctaaaaa 79

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: gb\_est1:\*

2: gb\_est2:\*

3: gb\_est3:\*

4: gb\_hc:\*

5: gb\_est4:\*

6: gb\_est5:\*

7: gb\_est6:\*

8: gb\_est7:\*

9: gb\_gsl1:\*

10: gb\_gsl2:\*

11: gb\_gsl3:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	76.4	96.7	479	5 BX108872	BF431778 nab50c04.
C 2	75.8	95.9	292	2 BF431778	BM053069 ie69h03.y
C 3	75.8	95.9	313	3 BM053069	CK903980 ie69h03.y
C 4	75.8	95.9	338	7 CK903979	CK903979 ie69h03.y
C 5	75.8	95.9	339	7 CK903979	AI927624 wo85a02.x
C 6	75.8	95.9	475	1 AI927624	BQ182988 UI-H-ED1-
C 7	75.8	95.9	633	3 BQ182988	CB241372 UI-CF-FNO
C 8	75.8	95.9	684	6 CB241372	CD364677 UI-H-F72-
C 9	75.8	95.9	691	6 CD364677	BF002331 7h22b08.x
C 10	75.8	95.9	716	2 BF002331	AA826896 nr86c10.s
C 11	74.8	94.7	109	1 AA826896	AW874230 hg85h05.x
C 12	74.8	94.7	180	1 AW874230	BF002304 7h02g04.x
C 13	74.8	94.7	272	2 BF002304	AA903087 ok46d04.s
C 14	74.8	94.7	310	1 AA903087	AI621135 tu49d01.x
C 15	74.8	94.7	332	1 AI621135	AI281889 qt82a11.x
C 16	74.8	94.7	376	1 AI281889	AW002382 wu61c03.x
C 17	74.8	94.7	417	1 AW002382	AW594044 hg35d12.x
C 18	74.8	94.7	418	1 AW594044	AW195936 xi85a06.x
C 19	74.8	94.7	420	1 AW195936	AW207187 UI-H-B11-
C 20	74.8	94.7	429	1 AW207187	AW057702 wx02c07.x
C 21	74.8	94.7	432	1 AW057702	AI218726 oo07a10.x
C 22	74.8	94.7	440	1 AI218726	

C 23	74.8	94.7	442	1 AI652267	AI652267 wb28d06.x
C 24	74.8	94.7	453	1 AI805092	AI805092 tu33g10.x
C 25	74.8	94.7	522	2 BF939675	BF939675 nac79g07.
C 26	74.8	94.7	687	5 BU689727	BU689727 UI-CF-FNO
C 27	74.8	94.7	1633	4 CR595956	CR595956 full-leng
C 28	74.8	94.7	1633	4 CR595956	AW081554 xc42g10.x
C 29	73.8	93.4	429	1 AW081554	AI979001 wr69a10.x
C 30	73.8	93.4	456	1 AI979001	BE677472 7d84g07.x
C 31	73	92.4	408	2 BE677472	AI433486 ti53b01.x
C 32	73	92.4	413	1 AI433486	AI433486 ti53b01.x
C 33	73	92.4	413	1 AI433494	CR594943 full-leng
C 34	72.6	91.9	1650	4 CR594943	AW014948 UI-H-B10-
C 35	72.6	91.9	519	1 AW014948	AW048392 NISC-g704
C 36	71.8	90.9	693	6 CB048392	BX378127 BX378127
C 37	71.6	90.6	322	5 BX378127	AW974325 EST386428
C 38	71.4	90.4	350	1 AW974325	AI277084 ql40h08.x
C 39	70.6	89.4	157	1 AI277084	AI886304 wm67a04.x
C 40	63.6	80.5	145	1 AI886304	AI480347 tm51c09.x
C 41	63.6	80.5	1064	5 BX335528	CO1591 HUMGS000859
C 42	63.4	80.3	209	5 CO1591	AX380143 BX380143
C 43	62.2	78.7	799	5 BX380143	AW058447 wx21a06.x
C 44	62	78.5	445	1 AW058447	CR614987 full-leng
C 45	58	73.4	1700	4 CR614987	

# ALIGNMENTS

RESULT 1  
BX108872/c  
LOCUS  
DEFINITION BX108872 NCI CGAP Col4 Homo sapiens CDNA clone IMAGp998A214818 ;  
IMAGE:1961756, mRNA sequence.  
ACCESSION BX108872  
VERSION BX108872.1 GI:27835474  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

479 bp mRNA linear EST 07-FEB-2003  
BX108872 NCI CGAP Col4 Homo sapiens CDNA clone IMAGp998A214818 ;  
IMAGE:1961756, mRNA sequence.

REFERENCE  
AUTHORS Ebert,L., Heil,O., Hennig,S., Neubert,P., Partsch,E., Peters,M.,  
Radelof,U., Schneider,D. and Korn,B.  
TITLE Human UnigeneSet - RZPD3  
JOURNAL Unpublished (2003)  
COMMENT Contact: Ina Rolfs  
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH  
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany  
RZPD; IMAGp998A214818.  
RZPDLJB; I.M.A.G.E. cDNA Clone Collection;  
Human UnigeneSet - RZPD3 (RZPDLJB No.972)  
http://www.rzpd.de/CloneCards/cgi-bin/showlib.pl.cgi/response?libNo=972 Contact: Ina Rolfs  
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH  
Heubnerweg 6, D-14059 Berlin, Germany  
Tel: +49 30 32639 101  
Fax: +49 30 32639 111  
www.rzpd.de

This clone is available royalty-free from RZPD;  
contact RZPD (clone@rzpd.de) for further information. Seq primer:  
M13u, Primer sequence: CGTTGTAAACGACGGCCAGT.

FEATURES  
Location/Qualifiers  
1..479  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGp998A214818 ; IMAGE:1961756"  
/tissue\_type="moderately-differentiated adenocarcinoma"  
/lab\_host="DH10B"  
/clone\_lib="NCI CGAP Col4"  
/note="Organ: colon; Vector: pCMV-SPORT6; Site\_1: SalI;  
Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dt.

Average insert size 1.7 kb. Life Technologies catalog #: 11531-019"

```

ORIGIN
Query Match          96.7%; Score 76.4; DB 5; Length 479;
Best Local Similarity 97.5%; Pred. No. 8e-13;
Matches 77; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCCCTTCAGTTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
    |||||
Db 102 CGGGAATCCCCCTTCAGTTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 43
    |||||

Qy 61 AACAAACCGACTCTAAAAA 79
    |||||
Db 42 AACAAACCGACTCNACAAA 24
    |||||

RESULT 2
BF431778/c          292 bp      mRNA      linear      EST 19-JAN-2001
LOCUS
DEFINITION
nab50c04.x1 Soares_NSF_F8_9W_OT_PA_P_S1 Homo sapiens cDNA clone
IMAGE:3269358 3', mRNA sequence.

ACCESSION
BF431778
VERSION
BF431778.1 GI:11443892
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 292)
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgap@remail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40UP from Gbco
High quality sequence stop: 272.
Location/Qualifiers
1..292
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3269358"
/lab_host="DH10B"
/clone_lib="Soares NSF_F8_9W_OT_PA_P_S1"
/notes="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not 1; Site_2: Eco RI,
Equal amounts of plasmid DNA from five normalized
libraries were mixed, and ss circles were made in vitro.
Following HAP purification, this DNA was used as tracer in
a subtractive hybridization reaction. The driver was
PCR-amplified cDNAs from pools of 5,000 clones made from
the same 5 libraries. The pools consisted of the following
libraries and clones: Soares NBHSF pool 1:
309384-310919, 323208-325895 Soares Nb2HP pool 1:
145032-147335, 147720-148103, 148872-149255, 15002 -
150407, 151176-152327 Soares Nb2HF8-9W pool 1:
758280-760583, 772104-774407 Soares NbHPA pool 1:
304776-306311, 320136-322823, 326280-326663 Soares NBHOT
pool 1: 723720-726407, 739080-740999 Subtraction by Bento
Soares and M. Fatima Bonaldo."

FEATURES
source
1..313
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5672308"
/sex="Both"
/tissue_type="Islets of Langerhans"
/dev_stage="Adult"
/lab_host="DH10B"
/clone_lib="Melton Normalized Human Islet 4 N4-HIS 1"
/notes="Organ: Pancreas; Vector: pSPORT1; Site 1: Not 1;
Site 2: Sal 1; Starting library constructed using
SuperScript Plasmid Library kit (Life Technologies). cDNA
made by oligo-dt priming. Size-selected by column
fractionation; average insert size 1.08 Kb. Library was
amplified once on solid support and plasmid DNA from
library was prepared. The library DNA was normalized by
method #4 from Bonaldo, Lennon, and Soares 1996 Genome
Research 6:791-806; 0.5 microgram single-stranded library
plasmid DNA was mixed with 5 micrograms PCR product
representing library inserts and hybridized to an EcoT of
20. Single-stranded (unhybridized) plasmids were isolated
by hydroxyapatite chromatography and used to make this
library."

ORIGIN
Query Match          95.9%; Score 75.8; DB 2; Length 292;
Best Local Similarity 97.5%; Pred. No. 1.1e-12;
Matches 77; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCCCTTCAGTTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
    |||||

```

```

Db 81 CGGGAATCCCCCTTCAGTTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 22

Qy 61 AACAAACCGACTCTAAAAA 79
    |||||
Db 21 AACAAACCGACTCACAAA 3
    |||||

RESULT 3
BM053069
LOCUS
DEFINITION
1e69H03.y3 Melton Normalized Human Islet 4 N4-HIS 1 Homo sapiens
cDNA clone IMAGE:5672308 5', mRNA sequence.

ACCESSION
BM053069
VERSION
BM053069.1 GI:16809019
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 313)
Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,
Lemishka, I., Searce, M., Brestelli, J., Gradwohl, G., Clifton, S.,
Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blistain, A.,
Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J.,
Cardenas, M., Gibbons, M., McCann, R., Cole, R., Tsagarisvili, R.,
Williams, T., Jackson, Y. and Bowers, Y.
Endocrine Pancreas Consortium
Unpublished (2000)
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Douglas Melton DNA sequencing by:
Washington University Genome Sequencing Center. For information on
obtaining a clone please contact: Juliana Brown
(brown@fas.harvard.edu) This sequence now available from the IMAGE
consortium, for clone orders contact: info@image.llnl.gov.

FEATURES
source
1..313
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5672308"
/sex="Both"
/tissue_type="Islets of Langerhans"
/dev_stage="Adult"
/lab_host="DH10B"
/clone_lib="Melton Normalized Human Islet 4 N4-HIS 1"
/notes="Organ: Pancreas; Vector: pSPORT1; Site 1: Not 1;
Site 2: Sal 1; Starting library constructed using
SuperScript Plasmid Library kit (Life Technologies). cDNA
made by oligo-dt priming. Size-selected by column
fractionation; average insert size 1.08 Kb. Library was
amplified once on solid support and plasmid DNA from
library was prepared. The library DNA was normalized by
method #4 from Bonaldo, Lennon, and Soares 1996 Genome
Research 6:791-806; 0.5 microgram single-stranded library
plasmid DNA was mixed with 5 micrograms PCR product
representing library inserts and hybridized to an EcoT of
20. Single-stranded (unhybridized) plasmids were isolated
by hydroxyapatite chromatography and used to make this
library."

ORIGIN
Query Match          95.9%; Score 75.8; DB 3; Length 313;
Best Local Similarity 97.5%; Pred. No. 1.2e-12;
Matches 77; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```



```

Qy 1 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
Db 229 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 288
Qy 61 AACAAACCGACTCTAATAA 79
Db 289 AACAAACCGACTCACAATA 307

RESULT 4
CK903980
LOCUS
DEFINITION ie69h03.y5 Melton Normalized Human Islet 4 N4-HIS 1 Homo sapiens
cDNA clone IMAGE:5672308 5', mRNA sequence.
CK903980
ACCESSION CK903980.1 GI:45365511
VERSION
KEYWORDS
SOURCE EST.
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1 (bases 1 to 338)
Melton,D., Meadows,A., Clifton,S., Hillier,L., Marra,M., Pape,D.,
Wylie,T., Martin,J., Blistain,A., Schmitt,A., Theising,B.,
Ritter,E., Ronko,I., Bennett,J., Cardenas,M., Gibbons,M.,
McCann,R., Cole,R., Tsagareishvili,R., Williams,T., Jackson,Y. and
Bowers,Y.
WashU-Harvard Pancreas EST Project
Unpublished (2000)
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
This read is a 5' RESEQUENCE of a previously sequenced pancreas
clone
This read has been verified (found to hit its original self in the
correct orientation)
Putative full length read
vector to vector length is
Seq primer: -40RP from Gibco.
Location/Qualifiers
1..338
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5672308"
/sex="Both"
/tissue_type="Islets of Langerhans"
/dev_stage="Adult"
/lab_host="DH10B"
/clone_lib="Melton Normalized Human Islet 4 N4-HIS 1"
/note="Organ: Pancreas; Vector: pSPORT1; Site_1: Not 1;
Site_2: Sal 1; Starting library constructed using
SuperScript Plasmid Library kit (Life Technologies). cDNA
made by oligo-dT priming. Size-selected by column
fractionation; average insert size 1.08 kb. Library was
amplified once on solid support and plasmid DNA from
library #4 from Bonaldo, Lennon, and Soares 1996 Genome
Research 6:791-806; 0.5 microgram single-stranded library
plasmid DNA was mixed with 5 micrograms PCR product
representing library inserts and hybridized to an Ecot of
20. Single-stranded (unhybridized) plasmids were isolated
by hydroxyapatite chromatography and used to make this
library."

FEATURES
source
1..338
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5672308"
/sex="Both"
/tissue_type="Islets of Langerhans"
/dev_stage="Adult"
/lab_host="DH10B"
/clone_lib="Melton Normalized Human Islet 4 N4-HIS 1"
/note="Organ: Pancreas; Vector: pSPORT1; Site_1: Not 1;
Site_2: Sal 1; Starting library constructed using
SuperScript Plasmid Library kit (Life Technologies). cDNA
made by oligo-dT priming. Size-selected by column
fractionation; average insert size 1.08 kb. Library was
amplified once on solid support and plasmid DNA from
library #4 from Bonaldo, Lennon, and Soares 1996 Genome
Research 6:791-806; 0.5 microgram single-stranded library
plasmid DNA was mixed with 5 micrograms PCR product
representing library inserts and hybridized to an Ecot of
20. Single-stranded (unhybridized) plasmids were isolated
by hydroxyapatite chromatography and used to make this
library."

ORIGIN
Query Match 95.9%; Score 75.8; DB 7; Length 338;

Best Local Similarity 97.5%; Pred. No. 1.2e-12;
Matches 77; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
Db 247 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 306
Qy 61 AACAAACCGACTCTAATAA 79
Db 307 AACAAACCGACTCACAATA 325

RESULT 5
CK903979/c
LOCUS
DEFINITION ie69h03.x5 Melton Normalized Human Islet 4 N4-HIS 1 Homo sapiens
cDNA clone IMAGE:5672308 3', mRNA sequence.
CK903979
ACCESSION CK903979.1 GI:45365510
VERSION
KEYWORDS
SOURCE EST.
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1 (bases 1 to 339)
Melton,D., Meadows,A., Clifton,S., Hillier,L., Marra,M., Pape,D.,
Wylie,T., Martin,J., Blistain,A., Schmitt,A., Theising,B.,
Ritter,E., Ronko,I., Bennett,J., Cardenas,M., Gibbons,M.,
McCann,R., Cole,R., Tsagareishvili,R., Williams,T., Jackson,Y. and
Bowers,Y.
WashU-Harvard Pancreas EST Project
Unpublished (2000)
Other ESTs: ie69h03.Y3
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
This read is a 3' RESEQUENCE of a previously sequenced pancreas
clone
This resequenced clone has not previously been sequenced on this
end, resequencing from this end represents new data
Seq primer: -400P from Gibco.
Location/Qualifiers
1..339
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5672308"
/sex="Both"
/tissue_type="Islets of Langerhans"
/dev_stage="Adult"
/lab_host="DH10B"
/clone_lib="Melton Normalized Human Islet 4 N4-HIS 1"
/note="Organ: Pancreas; Vector: pSPORT1; Site_1: Not 1;
Site_2: Sal 1; Starting library constructed using
SuperScript Plasmid Library kit (Life Technologies). cDNA
made by oligo-dT priming. Size-selected by column
fractionation; average insert size 1.08 kb. Library was
amplified once on solid support and plasmid DNA from
library #4 from Bonaldo, Lennon, and Soares 1996 Genome
Research 6:791-806; 0.5 microgram single-stranded library
plasmid DNA was mixed with 5 micrograms PCR product
representing library inserts and hybridized to an Ecot of
20. Single-stranded (unhybridized) plasmids were isolated
by hydroxyapatite chromatography and used to make this
library."

```

Query Match 95.9%; Score 75.8; DB 7; Length 339;  
 Best Local Similarity 97.5%; Pred. No. 1.2e-12;  
 Matches 77; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGAATCCCCCTTCAGTCTCTTTGAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60  
 Db 81 CGGAATCCCCCTTCAGTCTCTTTGAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 22  
 Qy 61 AACAAACCGACTCTAAAAA 79  
 Db 21 AACAAACCGACTCAAAA 3

RESULT 6  
 A1927624/c  
 LOCUS w085a02.x1 NCI\_CGAP\_Kid11 Homo sapiens cDNA clone IMAGE:2462090 3',  
 DEFINITION mRNA sequence.  
 ACCESSION A1927624  
 VERSION A1927624.1 GI:5663588  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Hominidae; Homo.  
 1 (bases 1 to 475)  
 NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 Unpublished (1997)  
 JOURNAL  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapsb-remail.nih.gov  
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.  
 Emmert-Buck, M.D., Ph.D.  
 cDNA Library Preparation: M. Bento Soares, Ph.D.  
 CDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone Distribution: NCI-CCAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 www-bio.llnl.gov/bbrp/image/image.html  
 Insert Length: 1810 Std Error: 0.00  
 Seq primer: -40UP from Gibco  
 High quality sequence stop: 456.

FEATURES  
 source  
 1..475  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:2462090"  
 /lab\_host="DH10B"  
 /clone\_lib="NCI\_CGAP Kid11"  
 /note="Organ: Kidney; Vector: pTV73D-Pac (Pharmacia) with  
 a modified polylinker; Site 1: Not I; Site 2: Eco RI;  
 plasmid DNA from the normalized library NCI\_CGAP Kid3 was  
 prepared, and ss circles were made in vitro. Following RAP  
 purification, this DNA was used as tracer in a subtractive  
 hybridization reaction. The driver was PCR-amplified cDNAs  
 from a pool of 5,000 clones made from the same library  
 (cloneIDs 1322376-1323911, 1456007-1456775, and  
 1500552-1502855). Subtraction by Bento Soares and M.  
 Fatima Bonaldo."

ORIGIN  
 Query Match 95.9%; Score 75.8; DB 1; Length 475;  
 Best Local Similarity 97.5%; Pred. No. 1.2e-12;  
 Matches 77; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 Qy 1 CGGAATCCCCCTTCAGTCTCTTTGAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60  
 Db 81 CGGAATCCCCCTTCAGTCTCTTTGAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 22

Qy 61 AACAAACCGACTCTAAAAA 79  
 Db 21 AACAAACCGACTCAAAA 3

RESULT 7  
 BQ182988/c  
 LOCUS BQ182988  
 DEFINITION  
 ACCESSION BQ182988  
 VERSION BQ182988.1  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Hominidae; Homo.  
 1 (bases 1 to 633)  
 NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 Unpublished (1997)  
 JOURNAL  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapsb-remail.nih.gov  
 Tissue Procurement: Dr. Jose Mercuende  
 CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
 CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov  
 The following repetitive elements were found in this cDNA  
 sequence: 276-341 >GC rich#Low\_complexity  
 Seq primer: M13 FORWARD  
 POLYA=Yes.

FEATURES  
 source  
 1..633  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5839650"  
 /issue\_type="Chondrosarcoma"  
 /dev\_stage="Adult"  
 /lab\_host="DH10B (Life Technologies)"  
 /clone\_lib="NCI\_CGAP ED1"  
 /note="Organ: Left Pubic Bone; Vector: pTV73-Pac  
 (Pharmacia) with a modified polylinker; Site 1: EcoR I;  
 Site 2: Not I; NCI\_CGAP ED1 is a normalized cDNA library  
 containing the following tissue(s): Chondrosarcoma cell  
 line CS5. The library was constructed according to  
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,  
 1996. First strand cDNA synthesis was primed with an  
 oligo-dT primer containing a Not I site. Double stranded  
 cDNA was ligated to an EcoR I adaptor, digested with Not  
 I, and cloned directionally into pTV73-Pac vector. The  
 oligonucleotide used to prime the synthesis of  
 first-strand cDNA contains a library tag sequence that is  
 located between the Not I site and the (dT)18 tail. The  
 sequence tag for this library is GCTCAAGGCT.  
 TAG TISSUE=chondrosarcoma  
 TAG LIB=UI-H-ED1  
 TAG\_SEQ=CGTCAAGGCT"

ORIGIN  
 Query Match 95.9%; Score 75.8; DB 3; Length 633;  
 Best Local Similarity 97.5%; Pred. No. 1.3e-12;  
 Matches 77; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGAATCCCCCTTCAGTCTCTTTGAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60  
 Db 91 CGGAATCCCCCTTCAGTCTCTTTGAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 32  
 Qy 61 AACAAACCGACTCTAAAAA 79

Db		31 AACAAACCGACTCAGAAA 13
RESULT 8		
LOCUS	CB241372	684 bp mRNA linear EST 12-FEB-2003
DEFINITION	UI-CF-FNO-afy-h-23-0-UI.s1 UI-CF-FNO Homo sapiens cDNA clone	
ACCESSION	CB241372	
VERSION	CB241372.1	GI:28363016
KEYWORDS	EST.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
REFERENCE	1 (bases 1 to 684)	
AUTHORS	Bonaldo,M.F., Lennon,G. and Soares,M.B.	
TITLE	Normalization and subtraction: two approaches to facilitate gene discovery	
JOURNAL	Genome Res. 6 (9), 791-806 (1996)	
PUBMED	889548	
COMMENT	Contact: McCray, PB University of Iowa 2024 University of Iowa Med Labs, Iowa City, IA 52242, USA Tel: 319 356 4866 Fax: 319 356 7171 Email: paul-mccray@uiowa.edu Tissue Procurement: Dr. M. J. Welsh, University of Iowa cDNA Library preparation: Dr. M. Bento Soares, University of Iowa CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa DNA Sequencing by: Dr. M. Bento Soares, University of Iowa Clone Distribution: Researchers may obtain clones from Research Genetics (www.regen.com) or from Open Biosystems (www.openbiosystems.com). The following repetitive elements were found in this cDNA sequence: 277-342, >GC rich#Low_complexity Seq primer: M13 FORWARD POLYA=Yes	
FEATURES	Location/Qualifiers	
source	1..684	
	/organism="Homo sapiens"	
	/mol_type="mRNA"	
	/db_xref="taxon:9606"	
	/clone="UI-CF-FNO-afy-h-23-0-UI"	
	/tissue_type="Human Lung Epithelial cells"	
	/lab_host="DH10B (Life Technologies) (T1 phage resistant)"	
	/clone_lib="UI-CF-FNO"	
	/note="Organ: Lung; Vector: pTZ19-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-CF-FNO is a subtracted cDNA library derived from two normalized Human lung epithelial cell libraries (ENI and DU1) The library was subtracted according to according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. For additional information, contact: bentso-soares@uiowa.edu	
	TAG_TISSUE=Human Lung Epithelial Cell Lines untreated LPS 6hr to LPS 24h	
	TAG LIB=UI-CF-FNO	
	TAG_SEQ=CTGTCAGGT	
ORIGIN		
Query Match	95.9%; Score 75.8; DB 6; Length 684;	
Best Local Similarity	97.5%; Pred. No. 1.3e-12;	
Matches	77; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	
QY	1 CGGGATCCCTTCAGTTCTTTGAAGAAGTTCATGCTCGAATATCTGAATGAAGAA 60	
Db		92 CGGGATCCCTTCAGTTCTTTGAAGAAGTTCATGCTCGAATATCTGAATGAAGAA 33
QY	61 AACAAACCGACTCTAAAAA 79	
Db		32 AACAAACCGACTCAGAAA 14
RESULT 9		
LOCUS	CD364677	691 bp mRNA linear EST 05-AUG-2004
DEFINITION	UI-H-FT2-bjm-i-14-0-UI.s1 NCI_CGAP FT2 Homo sapiens cDNA clone	
ACCESSION	CD364677	
VERSION	CD364677.1	GI:31148767
KEYWORDS	EST.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
REFERENCE	1 (bases 1 to 691)	
AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.	
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index	
JOURNAL	Unpublished (1997)	
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgabs-f@mail.nih.gov Tissue Procurement: Dr. Gary W. Hunninghake, U of I cDNA Library preparation: Dr. M. Bento Soares, University of Iowa CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa DNA Sequencing by: Dr. M. Bento Soares, University of Iowa Clone Distribution: Distribution information can be found at http://genome.uiowa.edu/distribution/cgap.html Seq primer: M13 FORWARD POLYA=Yes	
FEATURES	Location/Qualifiers	
source	1..691	
	/organism="Homo sapiens"	
	/mol_type="mRNA"	
	/db_xref="taxon:9606"	
	/clone="UI-H-FT2-bjm-i-14-0-UI"	
	/tissue_type="Alveolar Macrophage"	
	/dev_stage="Adult"	
	/lab_host="DH10B (Life Technologies)"	
	/clone_lib="NCI_CGAP_FT2"	
	/note="Organ: Lung; Vector: pTZ19-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP FT2 is a subtracted cDNA library constructed from a pool of 81 RNA samples from Alveolar Macrophages challenged with different treatments. The mRNA samples were a mixture of these conditions (times refer to incubations following isolation by bronchoalveolar lavage) (some normal donor macrophages were cultured in some of the conditions, other donor macrophages in different conditions). The mRNA samples were pooled for library construction. Control 0 hours; control 3 hours; control 24 hours; LPS 100 ng/ml, 3 hours; LPS 100 ng/ml, 24 hours; PMA 10 ng/ml, 3 hours; PMA 10 ng/ml, 24 hours; Klebsiella moi 10, 3 hours; Klebsiella moi 10, 24 hours; Staph aureus moi 10, 3 hours; Staph aureus moi 10, 24 hours; Adenoviral vector (Ad5 CMV eGFP), moi 500, 3 hours; Adenoviral vector (Ad5 CMV eGFP), moi 500, 24 hours; wt adenovirus moi 500, 3 hours; wt adenovirus moi 500, 24 hours; Ad vector + LPS 3 hours; Ad vector + LPS 24 hours; wt adenovirus + LPS 3 hours; wt adenovirus + LPS 24 hours. The library was subtracted according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. The tissue was provided by Dr. Gary W. Hunninghake of the University of Iowa.	
	TAG_TISSUE=Human Lung Alveolar Macrophage	
	TAG LIB=UI-H-FT2	
	TAG_SEQ=GCGCATGCCG	
ORIGIN		
Query Match	95.9%; Score 75.8; DB 6; Length 691;	
Best Local Similarity	97.5%; Pred. No. 1.3e-12;	
Matches	77; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	
QY	1 CGGGATCCCTTCAGTTCTTTGAAGAAGTTCATGCTCGAATATCTGAATGAAGAA 60	
Db		92 CGGGATCCCTTCAGTTCTTTGAAGAAGTTCATGCTCGAATATCTGAATGAAGAA 33
QY	61 AACAAACCGACTCTAAAAA 79	

```

Matches 77; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAATGAAGAA 60
  |||||||
Db 91 CGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAATGAAGAA 32
  |||||||

QY 61 AACAAACCGACTCTAAAAA 79
  |||||||
Db 31 AACAAACCGACTCACAANA 13

RESULT 10
BF002331/c
LOCUS
DEFINITION
7h22b08.x1 NCI CGAP Col6 Homo sapiens cDNA clone IMAGE:3316695 3',
similar to TR:000466 000466 K12 PROTEIN PRECURSOR. ; mRNA
sequence.
ACCESSION
BF002331
VERSION
BF002331.1 GI:10702606
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens

REFERENCE
1 (bases 1 to 716)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabs-rc@mail.nih.gov
Tissue Procurement: Ilan Kirsch, M.D., Michael R. Emmert-Buck,
M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL, send email to:
info@image.llnl.gov
Seq primer: -40UP from Glibco
High quality sequence stop: 462.

FEATURES
source
1..716
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3316695"
/tissue_types="colon tumor, RER+"
/lab_hosts="DH10B"
/clone_lib="NCI CGAP Col6"
/notes="Organ: colon; Vector: pT73D-Pac (Pharmacia) with a
modified polylinker; Site 1: Not 1; Site 2: Eco RI;
Plasmid DNA from the normalized library NCI CGAP Col6 was
prepared, and ss circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(cloneIDs 1057416-1061255, and 1144584-1145351).
Subtraction by Bento Soares and M. Fatima Bonaldo. "
```

## ORIGIN

```

Query Match 95.9%; Score 75.8; DB 2; Length 716;
Best Local Similarity 97.5%; Pred. No. 1.3e-12;
Matches 77; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAATGAAGAA 60
  |||||||
Db 80 CGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAATGAAGAA 21
  |||||||

QY 61 AACAAACCGACTCTAAAAA 79
  |||||||
Db 20 AACAAACCGACTCACAANA 2
```

## RESULT 11

```

AA826896/c
LOCUS
DEFINITION
nr86c10.s1 NCI CGAP_Pr24 Homo sapiens cDNA clone IMAGE:1174866 3',
mRNA sequence.
ACCESSION
AA826896
VERSION
AA826896.1 GI:2900893
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens

REFERENCE
1 (bases 1 to 109)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabs-rc@mail.nih.gov
Tissue Procurement: Suzanne L. Topalian, M.D., Robert K. Bright,
Ph.D.
cDNA Library Preparation: Stratagene, Inc.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 1235 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 104.
```

## FEATURES

```

source
1..109
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1174866"
/tissue_types="invasive tumor (cell line)"
/lab_hosts="SOLR (kanamycin resistant)"
/clone_lib="NCI CGAP Pr24"
/notes="Organ: prostate; Vector: Bluescript SK-; Site 1:
EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:
Oligo dT. Invasive prostate tumor cell line (HPV
immortalized). 5' adaptor sequence: 5' GAATTCGGCAGAG 3'
3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'
Average insert size: 1.0 kb."
```

## ORIGIN

```

Query Match 94.7%; Score 74.8; DB 1; Length 109;
Best Local Similarity 97.4%; Pred. No. 2.1e-12;
Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAATGAAGAA 60
  |||||||
Db 82 CGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAATGAAGAA 23
  |||||||

QY 61 AACAAACCGACTCTAAAAA 78
  |||||||
Db 22 AACAAACCGACTCACAANA 5
```

## RESULT 12

```

AW874230/c
LOCUS
DEFINITION
hq86h05.x1 NCI CGAP_Thy3 Homo sapiens cDNA clone IMAGE:3126297 3',
mRNA sequence.
ACCESSION
AW874230
VERSION
AW874230.1 GI:8008283
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 180)

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index

#### JOURNAL

##### COMMENT

Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)

Tissue Procurement: Klaus Kaserer, M.D., Chris Moskaluk, M.D.,  
Ph.D., Michael R. Emmert-Buck, M.D., P.H.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL

Cloning strategy: Washington University Genome Sequencing Center  
DNA distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL, send email to:  
[info@image.llnl.gov](mailto:info@image.llnl.gov)

Seq primer: -40UP from Gibco

High quality sequence stop: 168.

#### FEATURES

##### source

1..180

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:3126297"

/tissue\_type="follicular carcinoma"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="NCI CGAP Thy3"

/note="Organ: thyroid; Vector: pCMV-Sport6; Site 1: Not I;  
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.

Average insert size 1.4 kb. Library constructed by Life

Technologies."

Technologies."

Technologies."

#### ORIGIN

Query Match 94.7%; Score 74.8; DB 1; Length 180;

Best Local Similarity 97.4%; Pred. No. 2.2e-12;

Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCTTCAGTCTTTGAAAAGTTCATGACTCGAATATCTGAATGAAGAA 60

Db 85 CGGGAATCCCTTCAGTCTTTGAAAAGTTCATGACTCGAATATCTGAATGAAGAA 26

Qy 61 AACAAACCGACTCTAAAA 78

Db 25 AACAAACCGACTCTACAA 8

#### RESULT 13

BF002304/c

LOCUS

DEFINITION 702904.x1 NCI\_CGAP\_Col6 Homo sapiens cDNA clone IMAGE:3314838 3',

mRNA sequence.

ACCESSION

BF002304

KEYWORDS

EST.

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 272)

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)

Tissue Procurement: Ilan Kirsch, M.D., Michael R. Emmert-Buck,

M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D.

cDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL, send email to:

Seq primer: -40UP from Gibco.

#### FEATURES

##### source

1..272

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:3314838"

/tissue\_type="colon tumor, RER."

/lab\_host="DH10B"

/clone\_lib="NCI CGAP Col6"

/note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a

modified polylinker; Site 1: Not I; Site 2: Eco RI;

Plasmid DNA from the normalized library NCI CGAP Col0 was

prepared, and ss circles were made in vitro. Following HAP

purification, this DNA was used as tracer in a subtractive

hybridization reaction. The driver was PCR-amplified cDNAs

from a pool of 5,000 clones made from the same library

(cloneIDs 1057416-1061255, and 1144584-1145351).

Subtraction by Bento Soares and M. Fatima Bonaldo.

#### ORIGIN

Query Match 94.7%; Score 74.8; DB 2; Length 272;

Best Local Similarity 97.4%; Pred. No. 2.3e-12;

Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCTTCAGTCTTTGAAAAGTTCATGACTCGAATATCTGAATGAAGAA 60

Db 83 CGGGAATCCCTTCAGTCTTTGAAAAGTTCATGACTCGAATATCTGAATGAAGAA 24

Qy 61 AACAAACCGACTCTAAAA 78

Db 28 AACAAACCGACTCTACAA 6

#### RESULT 14

AA903087/c

LOCUS

DEFINITION Q46d04.s1 NCI\_CGAP\_Lei2 Homo sapiens cDNA clone

IMAGE:1516999 3',

ACCESSION

AA903087

VERSION

AA903087.1

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 310)

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)

unknown library type

Seq primer: -40ml3 fwd. ET from Amersham

High quality sequence stop: 295.

Location/Qualifiers

1..310

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:1516999"

/tissue\_type="leiomyosarcoma"

/lab\_host="DH10B"

/clone\_lib="NCI CGAP Lei2"

/note="Organ: soft tissue; Vector: pT7T3D-Pac (Pharmacia)

with a modified polylinker; Site 1: Not I; Site 2: Eco RI;

1st strand cDNA was primed with a Not I - oligo(dT) primer

[5'-AACTGGAAGAATTCGCGCGCAATCGTTTTTTTTTTTTTTTTTT-3'],

double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pTV3 vector. Library went through one round of normalization. Library constructed by Bento Soares and M. Fatima Bonaldo."

## ORIGIN

Query Match 94.7%; Score 74.8; DB 1; Length 310;  
 Best Local Similarity 97.4%; Pred. No. 2.3e-12;  
 Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGGGAATCCCCCTTCAGTCTTTTGAAGAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60  
 |||||  
 Db 90 CGGGAATCCCCCTTCAGTCTTTTGAAGAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 31  
 |||||

QY 61 AACAAACCGACTCTAAAA 78  
 |||||  
 Db 30 AACAAACCGACTCACAAA 13  
 |||||

## RESULT 15

AI621135/c  
 LOCUS  
 DEFINITION tu49d01.x1 NCI\_CGAP\_Pr28 Homo sapiens cDNA clone IMAGE:2254369 3',  
 mRNA sequence.  
 ACCESSION AI621135  
 VERSION AI621135.1 GI:4630261  
 KEYWORDS EST.  
 SOURCE  
 ORGANISM Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homnidae; Homo.

## REFERENCE

1 (bases 1 to 332)  
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 Unpublished (1997)

## COMMENT

Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs@mail.nih.gov](mailto:cgapbs@mail.nih.gov)  
 Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.  
 Emmert-Buck, M.D., Ph.D.  
 CDNA Library Preparation: M. Bento Soares, Ph.D.  
 CDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
[www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)  
 Insert Length: 582 Std Error: 0.00  
 Seq primer: -40UP from Gibco  
 High quality sequence stop: 320  
 POLYA=NO.

## FEATURES

## Source

## Location/Qualifiers

1. .332  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:2254369"  
 /sex="male"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 /clone\_lib="NCI\_CGAP\_Pr28"  
 /note="Organ: prostate; Vector: pTV3D-Pac (Pharmacia)  
 with a modified polylinker; Plasmid DNA from the  
 normalized library NCI CGAP Pr22 was prepared, and ss  
 circles were made in vitro. Following HAP purification,  
 this DNA was used as tracer in a subtractive hybridization  
 reaction. The driver was PCR-amplified cDNAs from a pool  
 of 5,000 clones made from the same library (clones  
 985608-986759, 1101192-1101959, and 1217928-1220615).  
 Subtraction by Bento Soares and M. Fatima Bonaldo."

## ORIGIN

Query Match 94.7%; Score 74.8; DB 1; Length 332;  
 Best Local Similarity 97.4%; Pred. No. 2.4e-12;  
 Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGGGAATCCCCCTTCAGTCTTTTGAAGAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60  
 |||||  
 Db 131 CGGGAATCCCCCTTCAGTCTTTTGAAGAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 72  
 |||||

QY 61 AACAAACCGACTCTAAAA 78  
 |||||  
 Db 71 AACAAACCGACTCACAAA 54  
 |||||

Search completed: December 7, 2005, 09:55:48  
 Job time : 904.825 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 10:04:41 ; Search time 38.8456 Seconds  
(without alignments)  
4072.612 Million cell updates/sec

Title: US-09-980-046B-4  
Perfect score: 89  
Sequence: 1 cgggaatcccccttcagtc.....aaaaaaaaaaaaaagaaaa 89

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA:\*

1: /cgn2\_6/ptodata/1/ina/1 COMB.seq:\*

2: /cgn2\_6/ptodata/1/ina/5 COMB.seq:\*

3: /cgn2\_6/ptodata/1/ina/6A COMB.seq:\*

4: /cgn2\_6/ptodata/1/ina/6B COMB.seq:\*

5: /cgn2\_6/ptodata/1/ina/H COMB.seq:\*

6: /cgn2\_6/ptodata/1/ina/PCTUS COMB.seq:\*

7: /cgn2\_6/ptodata/1/ina/PP COMB.seq:\*

8: /cgn2\_6/ptodata/1/ina/RE COMB.seq:\*

9: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	70	78.7	601	3	US-09-949-016-26334, A
C 2	70	78.7	601	3	US-09-949-016-105203, A
C 3	70	78.7	2000	3	US-09-997-165-3, Appli
4	70	78.7	2000	3	US-09-949-016-426, App
5	70	78.7	2001	3	US-09-949-016-2936, App
6	70	78.7	2180	2	US-08-755-559-2, Appli
7	70	78.7	2180	2	US-09-210-474-2, Appli
8	70	78.7	2180	3	US-09-539-774-2, Appli
9	70	78.7	16738	3	US-09-949-016-12168, A
10	70	78.7	16738	3	US-09-949-016-14678, A
11	43.6	49.0	177	3	US-09-621-976-8073, App
12	41	46.1	164	3	US-09-621-976-8070, App
13	41	45.1	166	3	US-09-621-976-8651, App
14	40.6	45.6	1477	3	US-09-585-173B-7, Appli
15	40.4	45.4	246	3	US-09-621-976-16288, A
16	40.4	45.4	259	3	US-09-621-976-16294, A
17	39.4	44.3	249	3	US-09-621-976-16291, A
18	38.6	43.4	249	3	US-09-621-976-16292, A
19	38.6	43.4	732	3	US-09-149-476-66, Appli
20	38.6	43.4	1117	3	US-09-247-373B-33, Appli
21	38.6	43.4	2091	3	US-09-813-818-1, Appli
22	38.6	43.4	2091	3	US-10-199-333-1, Appli
23	38.2	42.9	127	3	US-09-621-976-13933, A
24	38	42.7	159	3	US-09-621-976-8623, App

25	38	42.7	159	3	US-09-621-976-8655	Sequence 8655, App
26	37.8	42.5	1743	3	US-10-012-231A-235	Sequence 235, App
27	37.8	42.5	1743	3	US-10-015-389A-235	Sequence 235, App
28	37.8	42.5	1743	3	US-10-006-768A-235	Sequence 235, App
29	37.8	42.5	1743	3	US-10-015-671A-235	Sequence 235, App
30	37.8	42.5	1743	3	US-10-015-393A-235	Sequence 235, App
31	37.8	42.5	1743	3	US-10-011-833A-235	Sequence 235, App
32	37.8	42.5	1743	3	US-10-006-041A-235	Sequence 235, App
33	37.8	42.5	1743	3	US-10-012-064A-235	Sequence 235, App
34	37.4	42.0	266	3	US-09-621-976-16813, A	Sequence 16813, A
35	37.4	42.0	3275	3	US-09-370-838-151	Sequence 151, App
36	37.4	42.0	3275	3	US-09-854-133-151	Sequence 151, App
C 37	37.2	41.8	11050	3	US-09-949-016-14116	Sequence 14116, A
C 38	37	41.6	194	3	US-10-131-827-8633	Sequence 8633, App
39	36.8	41.3	1882	3	US-09-419-679-3	Sequence 3, Appli
40	36.6	41.1	1406	3	US-10-000-489-81	Sequence 81, Appli
41	36.6	41.1	1934	3	US-08-776-844-1	Sequence 1, Appli
42	36.6	41.1	1934	3	US-09-909-325-1	Sequence 1, Appli
43	36.6	41.1	1934	3	US-09-909-326-1	Sequence 1, Appli
44	36.6	41.1	5520	3	US-10-001-887-43	Sequence 43, Appli
C 45	36.4	40.9	226	3	US-10-131-827-8671	Sequence 8671, App

ALIGNMENTS

RESULT 1  
US-09-949-016-26334/C  
; Sequence 26334, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 26334  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-26334

Query Match 78.7%; Score 70; DB 3; Length 601;  
Best Local Similarity 93.6%; Pred. No. 1.6e-08;  
Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Oy 1 CGGGAATCCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60  
|||  
Db 129 CGGGAATCCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 70  
|||

Oy 61 AACAAACCAAAAAAAA 78  
|||  
Db 69 AACAAACCGACTCACAA 52  
|||

RESULT 2

US-09-949-016-105203/c  
; Sequence 105203, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307

```
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 105203
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-105203

Query Match          78.7%; Score 70; DB 3; Length 601;
Best Local Similarity 93.6%; Pred. No. 1.6e-08;
Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCCCTTCAGTCTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
Db 129 CGGGAATCCCCCTTCAGTCTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 70

Qy 61 AACAAACCAAAAAA 78
Db 69 AACAAACCGACTCACA 52

RESULT 3
US-09-997-165-3
; Sequence 3, Application US/09997165
; Patent No. 6762030
; GENERAL INFORMATION:
; APPLICANT: Lyman, Stewart D.
; APPLICANT: Farnlow, William C.
; TITLE OF INVENTION: LIGAND FOR CD7 AND METHODS OF USE THEREOF
; FILE REFERENCE: 2913-US
; CURRENT APPLICATION NUMBER: US/09/997,165
; CURRENT FILING DATE: 2001-11-27
; PRIOR APPLICATION NUMBER: PCT/US00/14612
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/136,450
; PRIOR FILING DATE: 1999-05-28
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 3
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (119)..(865)
US-09-997-165-3

Query Match          78.7%; Score 70; DB 3; Length 2000;
Best Local Similarity 93.6%; Pred. No. 1.7e-08;
Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCCCTTCAGTCTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
Db 1645 CGGGAATCCCCCTTCAGTCTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1704

Qy 61 AACAAACCAAAAAA 78
Db 1705 AACAAACCGACTCACA 1722

RESULT 4
US-09-949-016-426
; Sequence 426, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
```

```
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 426
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-426

Query Match          78.7%; Score 70; DB 3; Length 2000;
Best Local Similarity 93.6%; Pred. No. 1.7e-08;
Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCCCTTCAGTCTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
Db 1645 CGGGAATCCCCCTTCAGTCTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1704

Qy 61 AACAAACCAAAAAA 78
Db 1705 AACAAACCGACTCACA 1722

RESULT 5
US-09-949-016-2936
; Sequence 2936, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2936
; LENGTH: 2001
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-2936

Query Match          78.7%; Score 70; DB 3; Length 2001;
Best Local Similarity 93.6%; Pred. No. 1.7e-08;
Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCCCTTCAGTCTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
Db 1646 CGGGAATCCCCCTTCAGTCTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1705

Qy 61 AACAAACCAAAAAA 78
Db 1706 AACAAACCGACTCACA 1723

RESULT 6
US-08-755-559-2
```



; Sequence 2, Application US/0875559  
; Patent No. 5912142  
; GENERAL INFORMATION:  
; APPLICANT: KAUFMAN, RUSSEL E.  
; APPLICANT: SLENTZ-KESLER, KIMBERLY  
; TITLE OF INVENTION: GENE PRODUCT OVER EXPRESSED IN CANCER  
; TITLE OF INVENTION: CELLS  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHYP P.C.  
; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR  
; CITY: ARLINGTON  
; STATE: VIRGINIA  
; COUNTRY: U.S.A.  
; ZIP: 22201-4714  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/755,559  
; FILING DATE: 22-NOV-1996  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: WILSON, MARY J.  
; REGISTRATION NUMBER: 32,955  
; REFERENCE/DOCKET NUMBER: 1579-116  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 816-4000  
; TELEFAX: (703) 816-4100  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2180 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; US-08-755-559-2

Query Match 78.7%; Score 70; DB 2; Length 2180;  
Best Local Similarity 93.6%; Pred. No. 1.7e-08;  
Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60  
Db 1841 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1900

Qy 61 AACAAACCAAAAAA 78  
Db 1901 AACAAACCGACTCACA 1918

RESULT 7  
US-09-210-474-2  
; Sequence 2, Application US/09210474  
; Patent No. 6072034  
; GENERAL INFORMATION:  
; APPLICANT: KAUFMAN, RUSSEL E.  
; APPLICANT: SLENTZ-KESLER, KIMBERLY  
; TITLE OF INVENTION: GENE PRODUCT OVER EXPRESSED IN CANCER  
; TITLE OF INVENTION: CELLS  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHYP P.C.  
; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR  
; CITY: ARLINGTON  
; STATE: VIRGINIA  
; COUNTRY: U.S.A.  
; ZIP: 22201-4714  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/210,474  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/755,559  
; FILING DATE: 22-NOV-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: WILSON, MARY J.  
; REGISTRATION NUMBER: 32,955  
; REFERENCE/DOCKET NUMBER: 1579-116  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 816-4000  
; TELEFAX: (703) 816-4100  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2180 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; US-09-210-474-2

Query Match 78.7%; Score 70; DB 3; Length 2180;  
Best Local Similarity 93.6%; Pred. No. 1.7e-08;  
Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60  
Db 1841 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1900

Qy 61 AACAAACCAAAAAA 78  
Db 1901 AACAAACCGACTCACA 1918

RESULT 8  
US-09-539-774-2  
; Sequence 2, Application US/09539774  
; Patent No. 6350615  
; GENERAL INFORMATION:  
; APPLICANT: KAUFMAN, RUSSEL E.  
; APPLICANT: SLENTZ-KESLER, KIMBERLY  
; TITLE OF INVENTION: GENE PRODUCT OVER EXPRESSED IN CANCER  
; TITLE OF INVENTION: CELLS  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHYP P.C.  
; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR  
; CITY: ARLINGTON  
; STATE: VIRGINIA  
; COUNTRY: U.S.A.  
; ZIP: 22201-4714  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/539,774  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/210,474  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: WILSON, MARY J.  
; REGISTRATION NUMBER: 32,955  
; REFERENCE/DOCKET NUMBER: 1579-116  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 816-4000

```
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 2180 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: DNA (genomic)
US-09-539-774-2

Query Match      78.7%; Score 70; DB 3; Length 2180;
Best Local Similarity 93.6%; Pred. No. 1.7e-08;
Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCCCTTCAGTCTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
Db 1841 CGGGAATCCCCCTTCAGTCTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1900

Qy 61 AACAAACCAAAAAA 78
Db 1901 AACAAACCGACTCACA 1918

RESULT 9
US-09-949-016-12168
; Sequence 12168, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12168
; LENGTH: 16738
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12168

Query Match      78.7%; Score 70; DB 3; Length 16738;
Best Local Similarity 93.6%; Pred. No. 2e-08;
Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCCCTTCAGTCTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
Db 14383 CGGGAATCCCCCTTCAGTCTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 14442

Qy 61 AACAAACCAAAAAA 78
Db 14443 AACAAACCGACTCACA 14460

RESULT 10
US-09-949-016-14678
; Sequence 14678, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; SOFTWARE: Patent.pm

Query Match      49.0%; Score 43.6; DB 3; Length 177;
Best Local Similarity 69.0%; Pred. No. 0.04;
Matches 58; Conservative 1; Mismatches 25; Indels 0; Gaps 0;

Qy 6 ATCCCCCTTCAGTCTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAAACAA 65
Db 80 ATCTCTCTCCCATTTTGTGAATAGTTTCAATAGAAAAA 139

Qy 66 ACCAAAAA 89
Db 140 AAAAAA 163

RESULT 11
US-09-621-976-8073
; Sequence 8073, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 8073
; LENGTH: 177
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-8073

Query Match      49.0%; Score 43.6; DB 3; Length 177;
Best Local Similarity 69.0%; Pred. No. 0.04;
Matches 58; Conservative 1; Mismatches 25; Indels 0; Gaps 0;

Qy 6 ATCCCCCTTCAGTCTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAAACAA 65
Db 80 ATCTCTCTCCCATTTTGTGAATAGTTTCAATAGAAAAA 139

Qy 66 ACCAAAAA 89
Db 140 AAAAAA 163

RESULT 12
US-09-621-976-8070
; Sequence 8070, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
```

```
; SEQ ID NO 8070
; LENGTH: 164
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-8070

Query Match      46.1%; Score 41; DB 3; Length 164;
Best Local Similarity 69.1%; Pred. No. 0.17;
Matches 56; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

Qy 9 CCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAATGAAGAAAAACAACC 68
Db 82 CCTCCTCCATTTTGTGAATAGTTTCAATAGAAAAAAGAAAAAAGAAAAAAGAAAAA 141

Qy 69 AAAAAAAGAAAAAAGAAAAA 89
Db 142 AAAAAAAGAAAAAAGAAAAA 162

RESULT 13
US-09-621-976-8651
; Sequence 8651, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621.976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 8651
; LENGTH: 166
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-8651

Query Match      46.1%; Score 41; DB 3; Length 166;
Best Local Similarity 69.1%; Pred. No. 0.17;
Matches 56; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

Qy 9 CCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAATGAAGAAAAACAACC 68
Db 82 CCTCCTCCATTTTGTGAATAGTTTCAATAGAAAAAAGAAAAAAGAAAAAAGAAAAA 141

Qy 69 AAAAAAAGAAAAAAGAAAAA 89
Db 142 AAAAAAAGAAAAAAGAAAAA 162

RESULT 14
US-09-585-173B-7
; Sequence 7, Application US/09585173B
; Patent No. 6570063
; GENERAL INFORMATION:
; APPLICANT: Butler, Karlene
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Gutteridge, Steven
; APPLICANT: Maxwell, Carl
; TITLE OF INVENTION: Magnesium Chelataase
; FILE REFERENCE: BB1370 US NA
; CURRENT APPLICATION NUMBER: US/09/585,173B
; CURRENT FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: US 60/137,461
; PRIOR FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 7
; LENGTH: 1477
; TYPE: DNA
; ORGANISM: Oryza sativa
```

---

```
US-09-585-173B-7

Query Match      45.6%; Score 40.6; DB 3; Length 1477;
Best Local Similarity 69.6%; Pred. No. 0.25;
Matches 55; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

Qy 11 CCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAATGAAGAAAAACAACC 70
Db 1379 CGTACAGTCTTAAGTAATAATAACATTCATAATTTATGTGAAAAAAGAAAAAAGAAAAA 1438

Qy 71 AAAAAAAGAAAAAAGAAAAA 89
Db 1439 AAAAAAAGAAAAAAGAAAAA 1457

RESULT 15
US-09-621-976-16288
; Sequence 16288, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621.976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 16288
; LENGTH: 246
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-16288

Query Match      45.4%; Score 40.4; DB 3; Length 246;
Best Local Similarity 75.8%; Pred. No. 0.25;
Matches 50; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 24 GAAAAAGTTCATGACTCGAATATCTGAATGAAGAAAAACAACCACAAAAAAGAAAAA 83
Db 164 GAAAAAGCGGTAACTAATAGCTGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 223

Qy 84 GAAAAA 89
Db 224 AAAAAA 229

Search completed: December 6, 2005, 11:47:18
Job time : 38.8456 secs
```

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 10:04:41 ; Search time 128.321 Seconds  
(without alignments)  
4072.612 Million cell updates/sec

Title: US-09-980-046B-7  
Perfect score: 294

Sequence: 1 cggcgatcgcgcggttat.....agcgagcacctagacaaaaa 294

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA.\*

- 1: /cgn2\_6/ptodata/1/ina/1 COMB.seq.\*
- 2: /cgn2\_6/ptodata/1/ina/5 COMB.seq.\*
- 3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*
- 4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*
- 5: /cgn2\_6/ptodata/1/ina/H COMB.seq.\*
- 6: /cgn2\_6/ptodata/1/ina/PCTUS COMB.seq.\*
- 7: /cgn2\_6/ptodata/1/ina/PP COMB.seq.\*
- 8: /cgn2\_6/ptodata/1/ina/RE COMB.seq.\*
- 9: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	269.4	91.6	2369	2	US-08-883-534-4
2	269.4	91.6	2369	3	US-09-204-764-4
3	269.4	91.6	2921	3	US-09-949-016-1211
4	269.4	91.6	2921	3	US-09-949-016-1212
5	269.4	91.6	46492	3	US-09-949-016-12953
6	269.4	91.6	46492	3	US-09-949-016-12954
7	269	91.5	601	3	US-09-949-016-42279
8	269	91.5	601	3	US-09-949-016-42346
9	35.8	12.2	601	3	US-09-949-016-108017
10	35.8	12.2	32104	3	US-09-949-016-14722
11	31.4	10.7	12128	3	US-09-949-016-12587
12	31.4	10.7	12129	3	US-09-949-016-15713
13	30.8	10.5	1809	3	US-09-489-039A-5265
14	30.6	10.4	150	2	US-08-513-846-14
15	30.6	10.4	380	3	US-09-385-982-453
16	30.6	10.4	407	3	US-09-385-982-519
17	30.2	10.3	76563	3	US-09-949-016-17099
18	30.2	10.3	78269	3	US-09-949-016-12497
19	30	10.2	263	3	US-09-513-998C-34198
20	30	10.2	229354	3	US-09-705-400-64
21	30	10.2	462589	3	US-09-949-016-12900
22	30	10.2	476044	3	US-09-949-016-12412
23	29.6	10.1	2246	3	US-10-104-047-822
24	29.6	10.1	22118	3	US-09-815-981A-5

C 25	29.6	10.1	28257	3	US-09-949-016-13076
C 26	29.6	10.1	28438	3	US-09-820-790B-3
C 27	29.6	10.1	33112	3	US-10-429-873A-3
C 28	29.6	10.1	42693	3	US-09-949-016-17317
C 29	29.6	10.1	42693	3	US-09-949-016-17318
C 30	29.6	10.0	27727	3	US-09-949-016-15737
C 31	29.2	9.9	601	3	US-09-949-016-118858
C 32	29.2	9.9	601	3	US-09-949-016-118859
C 33	29.2	9.9	601	3	US-09-949-016-121206
C 34	29.2	9.9	601	3	US-09-949-016-121207
C 35	29.2	9.9	2007	3	US-09-620-312D-238
C 36	29.2	9.9	23218	3	US-09-949-016-11987
C 37	29.2	9.9	23219	3	US-09-949-016-13396
C 38	29.2	9.9	61158	3	US-09-949-016-15041
C 39	29.2	9.9	70563	3	US-09-949-016-16743
C 40	29.2	9.9	75431	3	US-09-949-016-15122
C 41	29	9.9	5538	3	US-09-949-016-14398
C 42	28.8	9.8	434	3	US-09-513-999C-10571
C 43	28.8	9.8	1641	3	US-09-375-932A-2
C 44	28.8	9.8	1641	3	US-10-156-562A-2
C 45	28.8	9.8	16998	3	US-09-676-610B-24

#### ALIGNMENTS

RESULT 1  
US-08-883-534-4  
; Sequence 4, Application US/08883534  
; Patent No. 5846777  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; APPLICANT: Lal, Preeti  
; APPLICANT: Corley, Neil C.  
; TITLE OF INVENTION: TWO NEW WD-40 PROTEINS  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/883,534  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0332 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2369 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: NEUTOMT01  
; CLONE: 1221143  
US-08-883-534-4

Sequence 13076, A  
Sequence 3, Appli  
Sequence 3, Appli  
Sequence 17317, A  
Sequence 17318, A  
Sequence 15737, A  
Sequence 118858,  
Sequence 118859,  
Sequence 121206,  
Sequence 121207,  
Sequence 238, App  
Sequence 13987, A  
Sequence 13396, A  
Sequence 15041, A  
Sequence 16743, A  
Sequence 15122, A  
Sequence 14398, A  
Sequence 10571, A  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 24, Appli

Query Match 91.6%; Score 269.4; DB 2; Length 2369;  
Best Local Similarity 99.6%; Pred. No. 2.8e-80;  
Matches 270; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 21 TCCCATGATGCTCTGTCAAGGAGTGGACAATCACTACTGAGGAGCCCAACCCCGCCT 80  
Db 1898 TCCCATGATGCTCTGTCAAGGAGTGGACAATCACTACTGAGGAGCCCAACCCCGCCT 1957

Qy 81 CTGATGGACCGAATCAGGACTAGAGTTTAACTGCGAGCGGAACATGTCTCTATT 140  
Db 1958 CTGATGGACCGAATCAGGACTAGAGTTTAACTGCGAGCGGAACATGTCTCTATT 2017

Qy 141 TCTGTGACGGCGCCCATGCCCCCACCACCAAGAGGAGGAGGCGCCAGTCATGAC 200  
Db 2018 TCTGTGACGGCGCCCATGCCCCCACCACCAAGAGGAGGAGGCGCCAGTCATGAC 2077

Qy 201 CCTCGTCTCTGACGGGTGTCTGTACAGCTTCTTCTGAAAGCTTTAGACAGTAACAGTTTG 260  
Db 2078 CCTCGTCTCTGACGGGTGTCTGTACAGCTTCTTCTGAAAGCTTTAGACAGTAACAGTTTG 2137

Qy 261 CACATGAAAAATAAAGCGAGCACCTAGACAA 291  
Db 2138 CACATGAAAAATAAAGCGAGCACCTAAACAA 2168

## RESULT 2

US-09-204-764-4  
; Sequence 4, Application US/09204764  
; Patent No. 6025464  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; APPLICANT: Lal, Preeti  
; APPLICANT: Corley, Neil C.  
; TITLE OF INVENTION: TWO NEW WD-40 PROTEINS  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/204,764  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/883,534  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PP-0332 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2369 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: NEUTGMT01  
; CLONE: 1221143  
US-09-204-764-4

Query Match 91.6%; Score 269.4; DB 3; Length 2369;  
Best Local Similarity 99.6%; Pred. No. 2.8e-80;  
Matches 270; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 21 TCCCATGATGCTCTGTCAAGGAGTGGACAATCACTACTGAGGAGCCCAACCCCGCCT 80  
Db 1898 TCCCATGATGCTCTGTCAAGGAGTGGACAATCACTACTGAGGAGCCCAACCCCGCCT 1957

Qy 81 CTGATGGACCGAATCAGGACTAGAGTTTAACTGCGAGCGGAACATGTCTCTATT 140  
Db 1958 CTGATGGACCGAATCAGGACTAGAGTTTAACTGCGAGCGGAACATGTCTCTATT 2017

Qy 141 TCTGTGACGGCGCCCATGCCCCCACCACCAAGAGGAGGAGGCGCCAGTCATGAC 200  
Db 2018 TCTGTGACGGCGCCCATGCCCCCACCACCAAGAGGAGGAGGCGCCAGTCATGAC 2077

Qy 201 CCTCGTCTCTGACGGGTGTCTGTACAGCTTCTTCTGAAAGCTTTAGACAGTAACAGTTTG 260  
Db 2078 CCTCGTCTCTGACGGGTGTCTGTACAGCTTCTTCTGAAAGCTTTAGACAGTAACAGTTTG 2137

Qy 261 CACATGAAAAATAAAGCGAGCACCTAGACAA 291  
Db 2138 CACATGAAAAATAAAGCGAGCACCTAAACAA 2168

## RESULT 3

US-09-949-016-1211  
; Sequence 1211, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1211  
; LENGTH: 2921  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-1211

Query Match 91.6%; Score 269.4; DB 3; Length 2921;  
Best Local Similarity 99.6%; Pred. No. 3e-80;  
Matches 270; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 21 TCCCATGATGCTCTGTCAAGGAGTGGACAATCACTACTGAGGAGCCCAACCCCGCCT 80  
Db 1982 TCCCATGATGCTCTGTCAAGGAGTGGACAATCACTACTGAGGAGCCCAACCCCGCCT 2041

Qy 81 CTGATGGACCGAATCAGGACTAGAGTTTAACTGCGAGCGGAACATGTCTCTATT 140  
Db 2042 CTGATGGACCGAATCAGGACTAGAGTTTAACTGCGAGCGGAACATGTCTCTATT 2101

Qy 141 TCTGTGACGGCGCCCATGCCCCCACCACCAAGAGGAGGAGGCGCCAGTCATGAC 200  
Db 2102 TCTGTGACGGCGCCCATGCCCCCACCACCAAGAGGAGGAGGCGCCAGTCATGAC 2161

Qy 201 CCTCGTCTCTGACGGGTGTCTGTACAGCTTCTTCTGAAAGCTTTAGACAGTAACAGTTTG 260  
Db 2162 CCTCGTCTCTGACGGGTGTCTGTACAGCTTCTTCTGAAAGCTTTAGACAGTAACAGTTTG 2221

Qy 261 CACATGAAAAATAAAGCGAGCACCTAGACAA 291  
Db 2138 CACATGAAAAATAAAGCGAGCACCTAAACAA 2168

Db 2222 CACATGAAAAATAAAGCGAGCACCTAAACAA 2252

## RESULT 4

US-09-949-016-1212  
; Sequence 1212, Application US/09949016  
; Patent No. 6812339

## GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1212  
; LENGTH: 2921  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-1212

Query Match 91.6%; Score 269.4; DB 3; Length 2921;  
Best Local Similarity 99.6%; Pred. No. 3e-80;  
Matches 270; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 21 TCCCATGATGCTCTGTCAAGGAGTGGACAATCACCTACTGAGGAGCCCAACCCCGCCT 80

Db 1982 TCCCATGATGCTCTGTCAAGGAGTGGACAATCACCTACTGAGGAGCCCAACCCCGCCT 2041

Qy 81 CTGGATGACCGAATCAGGAGTGTAACTGACGCGGAACATGTCATTTCTCTATT 140

Db 2042 CTGGATGACCGAATCAGGAGTGTAACTGACGCGGAACATGTCATTTCTCTATT 2101

Qy 141 TCTGTGACGCGCCCATGCCCCCACCACCAAGAGGCGAGGCGCCCGATCATGAC 200

Db 2102 TCTGTGACGCGCCCATGCCCCCACCACCAAGAGGCGAGGCGCCCGATCATGAC 2161

Qy 201 CCTCGTCTCTGACGGGTGTCTGTACACGTTTCTCTGAAAGCTTTAGACAGTAACAGTTTG 260

Db 2162 CCTCGTCTCTGACGGGTGTCTGTACACGTTTCTCTGAAAGCTTTAGACAGTAACAGTTTG 2221

Qy 261 CACATGAAAAATAAAGCGAGCACCTAGACAA 291

Db 2222 CACATGAAAAATAAAGCGAGCACCTAGACAA 2252

## RESULT 5

US-09-949-016-12953  
; Sequence 12953, Application US/09949016  
; Patent No. 6812339

## GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12953

; LENGTH: 46492

; TYPE: DNA

; ORGANISM: Human

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (1)...(46492)

; OTHER INFORMATION: n = A,T,C or G

US-09-949-016-12953

Query Match 91.6%; Score 269.4; DB 3; Length 46492;

Best Local Similarity 99.6%; Pred. No. 9.1e-80;  
Matches 270; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 21 TCCCATGATGCTCTGTCAAGGAGTGGACAATCACCTACTGAGGAGCCCAACCCCGCCT 80

Db 43553 TCCCATGATGCTCTGTCAAGGAGTGGACAATCACCTACTGAGGAGCCCAACCCCGCCT 43612

Qy 81 CTGGATGACCGAATCAGGAGTGTAACTGACGCGGAACATGTCATTTCTCTATT 140

Db 43613 CTGGATGACCGAATCAGGAGTGTAACTGACGCGGAACATGTCATTTCTCTATT 43672

Qy 141 TCTGTGACGCGCCCATGCCCCCACCACCAAGAGGCGAGGCGCCCGATCATGAC 200

Db 43673 TCTGTGACGCGCCCATGCCCCCACCACCAAGAGGCGAGGCGCCCGATCATGAC 43732

Qy 201 CCTCGTCTCTGACGGGTGTCTGTACACGTTTCTCTGAAAGCTTTAGACAGTAACAGTTTG 260

Db 43733 CCTCGTCTCTGACGGGTGTCTGTACACGTTTCTCTGAAAGCTTTAGACAGTAACAGTTTG 43792

Qy 261 CACATGAAAAATAAAGCGAGCACCTAGACAA 291

Db 43793 CACATGAAAAATAAAGCGAGCACCTAGACAA 43823

## RESULT 6

US-09-949-016-12954  
; Sequence 12954, Application US/09949016  
; Patent No. 6812339

## GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12954  
; LENGTH: 46492  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(46492)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-12954

Query Match 91.6%; Score 269.4; DB 3; Length 46492;

Best Local Similarity 99.6%; Pred. No. 9.1e-80;  
Matches 270; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 21 TCCCATGATGCTCTGTCAAGGAGTGGACAATCACCTACTGAGGAGCCCAACCCCGCCT 80

Db 43553 TCCCATGATGCTCTGTCAAGGAGTGGACAATCACCTACTGAGGAGCCCAACCCCGCCT 43612

Qy 81 CTGGATGACCGAATCAGGAGTGTAACTGACGCGGAACATGTCATTTCTCTATT 140

Db 43553 TCCCATGATGCTCTGTCAAGGAGTGGACAATCACCTACTGAGGAGCCCAACCCCGCCT 43612





Db 159 AGCAATTTCTCGGCTCAGCCTCTAGAGTTGCTGGAGTACAGGCACCAAGCCACTGTACC 218  
QY 125 ATGTCATTTCTCTATTCTGTGACGCGCCCGCCATGCCCGCCACCCACCAAGAGGCGAGG 184  
Db 219 TGGCTTCATTTGTTTGTGCTGAGCATCAATCTTTCCACTCCACTGTATGCTCCATA 278  
QY 185 AGGCGCCAGTCATGACCTCGTCTCTGCAGG 215  
Db 279 AGGCGCCAGACAGCTGTCAATATATGAGG 309

## RESULT 10

US-09-949-016-14722/c  
; Sequence 14722, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 14722  
; LENGTH: 32104  
; TYPE: DNA  
; ORGANISM: Human  
; US-09-949-016-14722

Query Match 12.2%; Score 35.8; DB 3; Length 32104;  
Best Local Similarity 52.3%; Pred. No. 0.36;  
Matches 79; Conservative 0; Mismatches 72; Indels 0; Gaps 0;  
QY 65 AGCGCCACCCGCTCTGGATGGACCGAATCAGGACTAGAGTTAACTGCGAGCGGAAC 124  
Db 5051 AGCAATTTCTCGCCTCAGCCTCTAGAGTTGCTGGAGTACAGGCACCAAGCCACTGTACC 4992  
QY 125 ATGTCATTTCTCTATTCTGTGACGCGCCCGCCATGCCCGCCACCAAGAGGCGAGG 184  
Db 4991 TGGCTTCATTTGTTTGTGCTGAGCATCAATCTTTCCACTCCACTGTATGCTCCATA 4932  
QY 185 AGGCGCCAGTCATGACCTCGTCTCTGCAGG 215  
Db 4931 AGGCGCCAGACAGCTGTCAATATGATGAGG 4901

## RESULT 11

US-09-949-016-12587/c  
; Sequence 12587, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12587

; LENGTH: 12128  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-12587

Query Match 10.7%; Score 31.4; DB 3; Length 12128;  
Best Local Similarity 54.9%; Pred. No. 7.4;  
Matches 62; Conservative 0; Mismatches 51; Indels 0; Gaps 0;  
QY 64 GAGCCCGACCCCGCCTCTGGATGGACCGAATCAGGACTAGAGTTTAACTGCGAGCGAA 123  
Db 2147 GACCGCGCGCCCTCTCCGCTCCCACTTCCGAGCGGTGAGTTCACTTTCCACACCC 2088  
QY 124 CATGTCATTTCTCTATTCTGTGACGCGCCCGCCATGCCCGCCACCCACCAAA 176  
Db 2087 CTTTTCAGCGCCACCCCTCCCTGACCGCGCTCCATTCGCGTCCCGCCACAGCCA 2035

## RESULT 12

US-09-949-016-15713/c  
; Sequence 15713, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 15713  
; LENGTH: 12129  
; TYPE: DNA  
; ORGANISM: Human  
; US-09-949-016-15713

Query Match 10.7%; Score 31.4; DB 3; Length 12129;  
Best Local Similarity 54.9%; Pred. No. 7.4;  
Matches 62; Conservative 0; Mismatches 51; Indels 0; Gaps 0;  
QY 64 GAGCCCGACCCCGCCTCTGGATGGACCGAATCAGGACTAGAGTTTAACTGCGAGCGAA 123  
Db 2147 GACCGCGCGCCCTCTCCGCTCCCACTTCCGAGCGGTGAGTTCACTTTCCACACCC 2088  
QY 124 CATGTCATTTCTCTATTCTGTGACGCGCCCGCCATGCCCGCCACCCACCAAA 176  
Db 2087 CTTTTCAGCGCCACCCCTCCCTGACCGCGCTCCATTCGCGTCCCGCCACAGCCA 2035

## RESULT 13

US-09-489-039A-5265/c  
; Sequence 5265, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; FILE REFERENCE: 2709-2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 5265  
; LENGTH: 1809  
; TYPE: DNA



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 21:43:57 ; Search time 1175.67 Seconds  
(without alignments)  
4099.008 Million cell updates/sec

Title: US-09-980-046B-8  
Perfect score: 103  
Sequence: 1 cggacagacggcagcagtc.....tgagcccttcctcacaaaaa 103

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_hic:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_est7:\*  
9: gb\_ges1:\*  
10: gb\_ges2:\*  
11: gb\_ges3:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	103	100.0	303	2	BG181484
2	101.4	98.4	180	1	A1990634
3	101.4	98.4	302	2	BG196487
4	101.4	98.4	304	2	BG203670
5	101.4	98.4	532	3	BM993129
6	101.4	98.4	538	1	A1141684
7	101.4	98.4	601	3	BQ004264
8	101.4	98.4	611	5	BX099653
9	101.4	98.4	670	6	CA310277
10	101.4	98.4	672	3	BM666896
11	101.4	98.4	672	5	BU628665
12	101.4	98.4	673	3	BM982024
13	101.4	98.4	687	6	CD742559
14	101.4	98.4	1874	4	BC025710
15	100.4	97.5	641	2	BF439942
16	100.4	97.5	690	2	BG054851
17	99.8	96.9	304	2	BG192962
18	99.8	96.9	638	5	BU731752
19	98.4	95.5	223	1	AW071613
20	98.4	95.5	1841	4	CR607140
21	98.2	95.3	304	2	BG194006
22	98.2	95.3	421	1	AA826047

C 23	98.2	95.3	480	1	AA977801
C 24	97	94.2	493	2	BE677186
C 25	96.8	94.0	457	8	N50984
C 26	96.6	93.8	251	1	AW084330
C 27	96.6	93.8	303	2	BG214210
C 28	96.4	93.6	321	1	AA252172
C 29	96.4	93.6	500	3	BP392226
C 30	95	92.2	244	1	AA931748
C 31	90.4	87.8	268	1	AI280735
C 32	90.4	87.8	419	2	BG686255
C 33	87.8	85.2	332	1	AA994205
C 34	86	83.5	622	3	BM999231
C 35	83.2	80.8	374	8	T64144
C 36	75.4	73.2	300	1	AI361098
C 37	75.4	73.2	417	1	AW014432
C 38	75.4	73.2	701	6	CD366471
C 39	74.4	72.2	295	1	AW517751
C 40	74.4	72.2	322	1	AW238793
C 41	74.4	72.2	345	3	BM126146
C 42	74.4	72.2	379	3	BM126439
C 43	74.4	72.2	484	3	BM679324
C 44	74.4	72.2	690	2	BF969846
C 45	74.4	72.2	706	3	BM679069

ALIGNMENTS

RESULT 1  
BG181484  
LOCUS RST332 Athersys RAGE Library Homo sapiens cdNA, mRNA linear EST 21-APR-2001  
DEFINITION BG181484  
ACCESSION BG181484.1 GI:13703171  
VERSION  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 303)  
AUTHORS Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R., Cain,S., Leventhal,C., Thornton,M., Ramachandran,R., Whittington,J., Lerner,L., Costanzo,D., McElligott,K., Booser,S., Mays,R., Smith,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K., Offenbacher,J., Danzig,J. and Ducar,M.  
TITLE Creation of genome-wide protein expression libraries using random activation of gene expression  
JOURNAL Nat. Biotechnol. 19 (5), 440-445 (2001)  
PUBMED 11329013  
COMMENT Contact: Scott J. Cain  
Athersys, Inc.  
3201 Carnegie Ave, Cleveland, OH 44115, USA  
Tel: 216 431 9900  
Fax: 216 361 9596  
Email: scai@atersys.com  
High quality sequence stop: 303.  
Location/Qualifiers  
1. 303  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/cell\_line="HT1080"  
/clone\_lib="Athersys RAGE Library"  
/note="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."

ORIGIN

Query Match 100.0%; Score 103; DB 2; Length 303;

Best Local Similarity 100.0%; Pred. No. 2.4e-23;  
Matches 103; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGACAGCGGACGAGTCCAGCTCGGTTCTCTCGGTTTATTCTGTTAGATGAA 60  
|||||  
Db 193 CGGACAGCGGACGAGTCCAGCTCGGTTCTCTCGGTTTATTCTGTTAGATGAA 252  
|||||  
QY 61 ATGTTTCCCAATAAATAGGGGCATGAGCCCTTCTCACAAAA 103  
|||||  
Db 253 ATGTTTCCCAATAAATAGGGGCATGAGCCCTTCTCACAAAA 295  
|||||

RESULT 2  
LOCUS AI990634/c 180 bp mRNA linear EST 08-SEP-1999  
DEFINITION W222F08.X1 NCI\_CGAP\_GC6 Homo sapiens cDNA clone IMAGE:2497959 3',  
mRNA sequence.  
ACCESSION AI990634  
VERSION AI990634.1 GI:5837515  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.

1 (bases 1 to 180)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index

Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapsb@mail.nih.gov](mailto:cgapsb@mail.nih.gov)

Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael  
R. Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima  
Bonaldo, Ph.D.

cDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
[www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)

Seq primer: -40UP from Gibco.

Location/Qualifiers  
1..180

/organism="Homo sapiens"  
/mol\_type="mRNA"

/db\_xref="taxon:9606"  
/clones="IMAGE:2497959"

/tissue\_types="pooled germ cell tumors"  
/lab\_hosts="DH10B"

/clone\_lib="NCI\_CGAP GC6"  
/notes="vector: p773D-Fac (Pharmacia) with a modified

polylinker; Site 1: Not 1; Site 2: Eco RI; Plasmid DNA  
from the normalized library NCI\_CGAP GC4 was prepared, and

ss circles were made in vitro. Following HAP purification,  
this DNA was used as tracer in a subtractive hybridization

reaction. The driver was PCR-amplified cDNAs from a pool  
of 5,000 clones made from the same library (Cloneda

1257096-1258631, 1469064-1470983, and 1475592-1476743).  
Subtraction by Bento Soares and M. Fatima Bonaldo.

Query Match 98.4%; Score 101.4; DB 1; Length 180;  
Best Local Similarity 99.0%; Pred. No. 7.3e-23;  
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGACAGCGGACGAGTCCAGCTCGGTTCTCTCGGTTTATTCTGTTAGATGAA 60  
|||||

Db 108 CGGACAGCGGACGAGTCCAGCTCGGTTCTCTCGGTTTATTCTGTTAGATGAA 49  
|||||

QY 61 ATGTTTCCCAATAAATAGGGGCATGAGCCCTTCTCACAAAA 103  
|||||

Db 48 ATGTTTCCCAATAAATAGGGGCATGAGCCCTTCTCACAAAA 6  
|||||

RESULT 3  
LOCUS BG196487

DEFINITION RST15710 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.

ACCESSION BG196487

VERSION BG196487.1 GI:13718174

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.

1 (bases 1 to 302)

Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,  
Cain,S., Leventhal,C., Thornton,M., Ramachandran,R.,  
Whittington,J., Lerner,L., Costanzo,D., McElligott,K., Booser,S.,  
Mays,R., Smith,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K.,  
Offenbacher,J., Danzig,J. and Ducar,M.

Creation of genome-wide protein expression libraries using random  
activation of gene expression

Nat. Biotechnol. 19 (5), 440-445 (2001)

11329013

Contact: Scott J. Cain

Athersys, Inc.

3201 Carnegie Ave, Cleveland, OH 44115, USA

Tel: 216 431 9900

Fax: 216 361 9596

Email: [scain@athersys.com](mailto:scain@athersys.com)

High quality sequence stop: 302.

Location/Qualifiers  
1..302

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/cell\_line="HT1080"

/clone\_lib="Athersys RAGE Library"

/notes="See 'Creation of Genome-wide Protein Expression  
Libraries using Random Activation of Gene Expression', the  
Nature Biotechnology, in press. Note that even though the  
cell type indicated is HT1080, since a random activation  
method was used, these sequence tags are not necessarily  
expressed in HT1080 under normal circumstances."

Query Match 98.4%; Score 101.4; DB 2; Length 302;  
Best Local Similarity 99.0%; Pred. No. 8.1e-23;  
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGACAGCGGACGAGTCCAGCTCGGTTCTCTCGGTTTATTCTGTTAGATGAA 60  
|||||

Db 191 CGGACAGCGGACGAGTCCAGCTCGGTTCTCTCGGTTTATTCTGTTAGATGAA 250  
|||||

QY 61 ATGTTTCCCAATAAATAGGGGCATGAGCCCTTCTCACAAAA 103  
|||||

Db 251 ATGTTTCCCAATAAATAGGGGCATGAGCCCTTCTCACAAAA 293  
|||||

RESULT 4  
LOCUS BG203670

DEFINITION RST23058 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.

ACCESSION BG203670

VERSION BG203670.1 GI:13725357

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.

1 (bases 1 to 304)

Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,

Cain, S., Leventhal, C., Thornton, M., Ramchandran, R., Whittington, J., Lerner, L., Costanzo, D., McElligott, K., Booser, S., Mays, R., Smith, E., Veloso, N., Klika, A., Hess, J., Cothren, K., Lo, K., Offenbacher, J., Danzig, J., and Ducar, M.  
 Creation of genome-wide protein expression libraries using random activation of gene expression  
 Nat. Biotechnol. 19 (5), 440-445 (2001)  
 11329013  
 COMMENT  
 Contact: Scott J. Cain  
 Athersys, Inc. Cleveland, OH 44115, USA  
 3201 Carnegie Ave, Cleveland, OH 44115, USA  
 Tel: 216 431 9900  
 Fax: 216 361 9596  
 Email: scain@atersys.com  
 High quality sequence stop: 304.  
 Location/Qualifiers  
 1..304

# FEATURES

Location/Qualifiers  
 1..304  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /cell\_line="HT1080"  
 /clone\_lib="Athersys RAGE Library"  
 /note="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."

## ORIGIN

Query Match 98.4%; Score 101.4; DB 2; Length 304;  
 Best Local Similarity 99.0%; Pred. No. 8.1e-23;  
 Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 CGGACAGCGCAGCAGTCCAGCTCTGGTTTCCTTCGTTTATCTGTAGATGAA 60  
 DB 193 CGGACAGCGCAGCAGTCCAGCTCTGGTTTCCTTCGTTTATCTGTAGATGAA 252  
 QY 61 ATGGTCCCAATAATAGGGCGATGAGCCCTTCTCAGAAAA 103  
 DB 253 ATGGTCCCAATAATAGGGCGATGAGCCCTTCTCAGAAAA 295

RESULT 5  
 BM993129/c  
 LOCUS  
 DEFINITION  
 IMAGE:5866052 3', mRNA sequence.  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homidae; Homo.  
 1 (bases 1 to 532)  
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 Unpublished (1997)  
 COMMENT  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbbs@mail.nih.gov  
 Tissue Procurement: Dr. Jose Mercuende  
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov  
 Seq primer: M13 FORWARD  
 POLYA=Yes.

# FEATURES

Location/Qualifiers  
 1..532

/organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5866052"  
 /tissue\_type="Metastatic Chondrosarcoma"  
 /dev\_stage="Adult"  
 /lab\_host="DH10B (Life Technologies)"  
 /clone\_lib="NCI CGAP DT0"  
 /notes="Organ; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site 1: Ecor I; Site 2: Not I;  
 NCI CGAP DT0 is a cDNA library containing the following tissue(s): Metastatic Chondrosarcoma in Lung. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is  
 AACTGTCGG.  
 TAG TISSUE=lung metastatic chondrosarcoma  
 TAG LIB=UI-H-DT0  
 TAG\_SEQ=AACTGTCGG"

## ORIGIN

Query Match 98.4%; Score 101.4; DB 3; Length 532;  
 Best Local Similarity 99.0%; Pred. No. 9.1e-23;  
 Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 CGGACAGCGCAGCAGTCCAGCTCTGGTTTCCTTCGTTTATCTGTAGATGAA 60  
 DB 114 CGGACAGCGCAGCAGTCCAGCTCTGGTTTCCTTCGTTTATCTGTAGATGAA 55  
 QY 61 ATGGTCCCAATAATAGGGCGATGAGCCCTTCTCAGAAAA 103  
 DB 54 ATGGTCCCAATAATAGGGCGATGAGCCCTTCTCAGAAAA 12

## RESULT 6

AI141684/c  
 LOCUS

DEFINITION

OT:08405.x1 NCI CGAP GC3 Homo sapiens cDNA clone IMAGE:1614249 3', similar to TR:009014 009014 PEPTIDE/HISTIDINE TRANSPORTER. ;, mRNA

sequence.

AI141684

VERSION

AI141684.1 GI:3649141

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Homidae; Homo.

1 (bases 1 to 538)

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgapbbs@mail.nih.gov

Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael

Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D.

cDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone Distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html

Seq primer: -40ml3 fwd. EF from Amersham

High quality sequence stop: 456.

Location/Qualifiers

1..538

source

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1614249"
/tissue_type="pooled germ cell tumors"
/lab_host="DH10B"
/clone_lib="NCI CGAP GC3"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from 3 pooled
germ cell tumors, and was then primed with a Not I -
oligo(dT) primer. Double-stranded cDNA was ligated to Eco
RI adaptors (Pharmacia), digested with Not I and cloned
into the Not I and Eco RI sites of the modified pT7T3
vector library is not normalized. Library was
constructed by Bento Soares and M. Fatima Bonaldo. "

ORIGIN
Query Match 98.4%; Score 101.4; DB 1; Length 538;
Best Local Similarity 99.0%; Pred. No. 9.1e-23;
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACACAGCGGAGCAGTCCAGCTCTGGTTTCCTTCGTTTATTCGTTAGATGAA 60
Db 113 CGACACAGCGGAGCAGTCCAGCTCTGGTTTCCTTCGTTTATTCGTTAGATGAA 54

QY 61 ATGTTCCCATTAATAAGGGCGATGAGCCCTTCTCACAATAA 103
Db 53 ATGTTCCCATTAATAAGGGCGATGAGCCCTTCTCACAATAA 11

RESULT 7
BQ004264/c
LOCUS
DEFINITION
UI-H-E10-ayn-o-17-0-UI.s1 NCI_CGAP_E10 Homo sapiens cDNA clone
IMAGE:5841256 3', mRNA sequence.
BQ004264.1 GI:19729164
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 601)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Seq primer: M13 FORWARD
PolyA=yes.
FEATURES
Location/Qualifiers
1..601
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5841256"
/tissue_type="Chondrosarcoma"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP E10"
/note="Organ: Left Pelvis; Vector: pT7T3-Pac (Pharmacia)
with a modified polylinker; Site 1: EcoR I; Site 2: Not I;
NCI_CGAP E10 is a cDNA library containing the following
tissue(s): Chondrosarcoma. The library was constructed
according to Bonaldo, Lennon and Soares, Genome Research,

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1614249"
/tissue_type="pooled germ cell tumors"
/lab_host="DH10B"
/clone_lib="NCI CGAP GC3"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from 3 pooled
germ cell tumors, and was then primed with a Not I -
oligo(dT) primer. Double-stranded cDNA was ligated to Eco
RI adaptors (Pharmacia), digested with Not I and cloned
into the Not I and Eco RI sites of the modified pT7T3
vector library is not normalized. Library was
constructed by Bento Soares and M. Fatima Bonaldo. "

ORIGIN
Query Match 98.4%; Score 101.4; DB 3; Length 601;
Best Local Similarity 99.0%; Pred. No. 9.3e-23;
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACACAGCGGAGCAGTCCAGCTCTGGTTTCCTTCGTTTATTCGTTAGATGAA 60
Db 114 CGACACAGCGGAGCAGTCCAGCTCTGGTTTCCTTCGTTTATTCGTTAGATGAA 55

QY 61 ATGTTCCCATTAATAAGGGCGATGAGCCCTTCTCACAATAA 103
Db 54 ATGTTCCCATTAATAAGGGCGATGAGCCCTTCTCACAATAA 12

RESULT 8
BX099653
LOCUS
DEFINITION
611 bp mRNA linear EST 06-FEB-2003
BX099653 Soares fetal liver spleen INFLS Homo sapiens cDNA clone
IMAGE:998115131 ; IMAGE:127958, mRNA sequence.
BX099653.1 GI:27830132
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 611)
Ebert, L., Heil, O., Hennig, S., Neubert, P., Partsch, E., Peters, M.,
Radloff, U., Schneider, D. and Korn, B.
Human Unigeneset - RZPD3
Unpublished (2003)
Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
RZPD; IMAGE:998115131.
RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
http://www.rzpd.de/CloneCards/cgi-
bin/showLib.pl.cgi/responsefile=972 Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 101
Fax: +49 30 32639 111
www.rzpd.de
This clone is available royalty-free from RZPD;
contact RZPD (clone@rzpd.de) for further information. Seq primer:
M13r, Primer sequence: TTTCACAGGAACAGCTATGAC.
FEATURES
Location/Qualifiers
1..611
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:998115131 ; IMAGE:127958"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares fetal liver spleen INFLS"
/note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
with a modified polylinker; Site 1: Pac I; Site 2: Eco RI;
1st strand cDNA was primed with a Pac I - oligo(dT) primer

```



```

/db xref="taxon:9606"
/clone="UI-E-CL1-aer-j-23-0-UI"
/tissue_type="human retina"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-CL1"
/notes="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-E-CL1 is a normalized cDNA library containing the
following tissue(s): retina. The library was constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into pT7T3-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tag for this library is CCGCG. This library was
created for the program, Gene Discovery in the Visual
System, supported by National Eye Institute (NEI).
TAG_TISSUE=human retina
TAG_LIB=UI-E-CL1
TAG_SEQ=CCGCG"

```

```

ORIGIN
Query Match      98.4%; Score 101.4; DB 3; Length 672;
Best Local Similarity 99.0%; Pred. No. 9.5e-23;
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGACAGCGCAGCAGTCCAGCTCTGGTTCTCTCGGTTTATTCGTTAGATGAA 60
    |||||
Db 117 CGGACAGCGCAGCAGTCCAGCTCTGGTTCTCTCGGTTTATTCGTTAGATGAA 58
    |||||

QY 61 ATGTTCCCATTAATAAGGGGCATGAGCCCTTCCTCAGAAAAA 103
    |||||
Db 57 ATGTTCCCATTAATAAGGGGCATGAGCCCTTCCTCAGAAAAA 15
    |||||

```

```

RESULT 11
BU628665/c
LOCUS
DEFINITION
UI-H-FG0-bdf-1-04-0-UI.s1 NCI CGAP_EN1_2 Homo sapiens cDNA clone
UI-H-FG0-bdf-1-04-0-UI 3', mRNA sequence.
ACCESSION
BU628665
VERSION
BU628665.1 GI:23294879
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1 (bases 1 to 672)
NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL
Unpublished (1997)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapb@mail.nih.gov
Tissue Procurement: James Martin
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone Distribution Information can be obtained
from Dr. M. Bento Soares, bent-soares@uiowa.edu
Seq primer: M13 FORWARD
POLYA=Yes.
Location/Qualifiers
1. .672
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-FG0-bdf-1-04-0-UI"

```

```

FEATURES
source
1. .672
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-FG0-bdf-1-04-0-UI"

```

```

/tissue_type="Enchondroma cell line"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP_EN1_2"
/notes="Organ: Bone; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
NCI CGAP_EN1_2 is a cDNA library containing the following
tissue(s): Enchondroma cell line (2 cell lines). The
library was constructed according to Bonaldo, Lennon and
Soares, Genome Research, 6:791-806, 1996. First strand
cDNA synthesis was primed with an oligo-dT primer
containing a Not I site. Double stranded cDNA was ligated
to an EcoR I adaptor, digested with Not I, and cloned
directionally into pT7T3-Pac vector. The oligonucleotide
used to prime the synthesis of first-strand cDNA contains
a library tag sequence that is located between the Not I
site and the (dT)18 tail. The sequence tag for this
library is CCGTCACTC. The cell lines was provided by Dr
James Martin from University of Iowa.
TAG_TISSUE=Enchondroma cell line (Mix of EN1 and EN2)
TAG_LIB=UI-H-FG0
TAG_SEQ=CGGTCACTC"

```

```

ORIGIN
Query Match      98.4%; Score 101.4; DB 5; Length 672;
Best Local Similarity 99.0%; Pred. No. 9.5e-23;
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGACAGCGCAGCAGTCCAGCTCTGGTTCTCTCGGTTTATTCGTTAGATGAA 60
    |||||
Db 114 CGGACAGCGCAGCAGTCCAGCTCTGGTTCTCTCGGTTTATTCGTTAGATGAA 55
    |||||

QY 61 ATGTTCCCATTAATAAGGGGCATGAGCCCTTCCTCAGAAAAA 103
    |||||
Db 54 ATGTTCCCATTAATAAGGGGCATGAGCCCTTCCTCAGAAAAA 12
    |||||

```

```

RESULT 12
BM982024/c
LOCUS
DEFINITION
UI-CF-EN1-adg-j-16-0-UI.s1 UI-CF-EN1 Homo sapiens cDNA clone
UI-CF-EN1-adg-j-16-0-UI 3', mRNA sequence.
ACCESSION
BM982024
VERSION
BM982024.1 GI:19605106
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1 (bases 1 to 673)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Tissue Procurement: Dr. M. J. Welsh, University of Iowa
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com) or from Open Biosystems
(www.openbiosystems.com).
Seq primer: M13 FORWARD
POLYA=Yes.
Location/Qualifiers
BM982024
UI-CF-EN1-adg-j-16-0-UI.s1 mRNA linear EST 21-FEB-2003
UI-CF-EN1-adg-j-16-0-UI 3', mRNA sequence.
ACCESSION
BM982024
VERSION
BM982024.1 GI:19605106
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1 (bases 1 to 673)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Tissue Procurement: Dr. M. J. Welsh, University of Iowa
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com) or from Open Biosystems
(www.openbiosystems.com).
Seq primer: M13 FORWARD
POLYA=Yes.
Location/Qualifiers

```

```

COMMENT
Contact: McCray, PB
McCray Lab
University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel: 319 356 4866
Fax: 319 356 7171
Email: paul-mccray@uiowa.edu
Tissue Procurement: Dr. M. J. Welsh, University of Iowa
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com) or from Open Biosystems
(www.openbiosystems.com).
Seq primer: M13 FORWARD
POLYA=Yes.
Location/Qualifiers

```



```
source
1. .673
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-CF-EN1-adg-j-16-0-UI"
/tissue_type="Primary Lung Cystic Fibrosis Epithelial Cells"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="UI-CF-EN1"
/notes="Organ: Lung; Vector: p77T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-CF-EN1 is a normalized cDNA library containing the following tissue(s): Primary Lung Cystic Fibrosis Epithelial Cells. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into p77T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (d)18 tail. The sequence tag for this library is CTGCTCAGGT. TAG TISSUE=Human Lung Epithelial Cell Lines untreated LPS 6hr to LPS 24h TAG LIB=UI-CF-EN1 TAG_SEQ=CTGCTCAGGT"

ORIGIN
Query Match 98.4%; Score 101.4; DB 3; Length 673;
Best Local Similarity 99.0%; Pred. No. 9.5e-23;
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGACAGCGCAGCAGTCCAGCTCTGGTTCTCTCGGTTTATTCGTAGATGAA 60
Db 117 CGGACAGCGCAGCAGTCCAGCTCTGGTTCTCTCGGTTTATTCGTAGATGAA 58

Qy 61 ATGGTTCCCATTAATAAGGGCGATGAGCCCTTCTCACAAAAA 103
|||||
Db 57 ATGGTTCCCATTAATAAGGGCGATGAGCCCTTCTCACAAAAA 15

RESULT 13
CD742559/c 687 bp mRNA linear EST 05-AUG-2004
LOCUS UI-H-FT2-bj1-e-08-0-UI.s1 NCI CGAP FT2 Homo sapiens cDNA clone
DEFINITION UI-H-FT2-bj1-e-08-0-UI 3', mRNA sequence.
ACCESSION CD742559
VERSION CD742559.1 GI:32293409
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 687)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Gary W. Hunninghake, U of I
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/cgap.html
Seq primer: M13 FORWARD
```

```
FEATURES
source
POLYA=Yes.
1. .687
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-FT2-bj1-e-08-0-UI"
/tissue_type="Alveolar Macrophage"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP FT2"
/notes="Organ: Lung; Vector: p77T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP FT2 is a subtracted cDNA library constructed from a pool of 81 RNA samples from Alveolar Macrophages challenged with different treatments. The mRNA samples were a mixture of these conditions (times refer to incubations following isolation by bronchoalveolar lavage) (some normal donor macrophages were cultured in some of the conditions, other donor macrophages in different conditions). The mRNA samples were pooled for library construction. Control 0 hours; control 3 hours; control 24 hours; LPS 100 ng/ml, 3 hours; LPS 100 ng/ml, 24 hours; PMA 10 ng/ml, 3 hours; PMA 10 ng/ml, 24 hours; Klebsiella aureus moi 10, 3 hours; Staph aureus moi 10, 24 hours; Staph aureus moi 10, 3 hours; Ad5 CMV eGFP, moi 500, 3 hours; Adenoviral vector (Ad5 CMV eGFP), moi 500, 24 hours; Adenovirus moi 500, 3 hours; wt adenovirus moi 500, 24 hours; Ad vector + LPS 3 hours; Ad vector + LPS 24 hours; wt adenovirus + LPS 3 hours; wt adenovirus + LPS 24 hours. The library was subtracted according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. The tissue was provided by Dr. Gary W. Hunninghake of the University of Iowa. TAG TISSUE=Human Lung Alveolar Macrophage TAG LIB=UI-H-FT2 TAG_SEQ=GGCCATGCGG"

ORIGIN
Query Match 98.4%; Score 101.4; DB 6; Length 687;
Best Local Similarity 99.0%; Pred. No. 9.5e-23;
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGACAGCGCAGCAGTCCAGCTCTGGTTCTCTCGGTTTATTCGTAGATGAA 60
|||||
Db 111 CGGACAGCGCAGCAGTCCAGCTCTGGTTCTCTCGGTTTATTCGTAGATGAA 52

Qy 61 ATGGTTCCCATTAATAAGGGCGATGAGCCCTTCTCACAAAAA 103
|||||
Db 51 ATGGTTCCCATTAATAAGGGCGATGAGCCCTTCTCACAAAAA 9

RESULT 14
BC025710 1874 bp mRNA linear HTC 19-NOV-2003
LOCUS BC025710
DEFINITION Homo sapiens solute carrier family 15, member 3, mRNA (cDNA clone IMAGE5224878), containing frame-shift errors.
ACCESSION BC025710
VERSION BC025710.1 GI:19343588
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 1874)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
```

Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.D., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodrigues, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalios, D.E., Schnerker, A., Schein, J.E., Jones, S.J., and Marra, M.A.

**TITLE**  
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

**JOURNAL**  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

**PUBMED**  
12477932

**REFERENCE**  
2 (bases 1 to 1874)

**AUTHORS**  
Strausberg, R.

**TITLE**  
Direct Submission

**JOURNAL**  
Submitted (06-MAR-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

**REMARK**  
NTH-MGC Project URL: <http://mgc.nci.nih.gov>

**COMMENT**  
Contact: MGC help desk  
Email: [cgapbs@mail.nih.gov](mailto:cgapbs@mail.nih.gov)  
Tissue Procurement: Life Technologies, Inc.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland  
Web site: <http://www.nisc.nih.gov/>  
Contact: [nisc\\_mgc@nri.nih.gov](mailto:nisc_mgc@nri.nih.gov)  
Akhter, N., Ayte, K., Becketrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Grant, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Masiello, C., Maskeri, B., Masrrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W., Turgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAC Plate: 49 Row: m Column: 19  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 7706116  
This clone has the following problem: frame shifted.

**FEATURES**  
source  
1..1874  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IWAGS:5224878"  
/tissue\_type="Pancreas, Spleen, adult pooled"  
/clone\_lib="NIH\_MGC\_120"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"

**ORIGIN**  
Query Match 98.4%; Score 101.4; DB 4; Length 1874;  
Best Local Similarity 99.0%; Pred. No. 1.2e-21;  
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGACAGCGGCGAGCTCCAGCTCTGGTTTCCTTCGTTTATCTGTTAGATGAA 60  
|||||  
Db 1707 CGGACAGCGGCGAGCTCCAGCTCTGGTTTCCTTCGTTTATCTGTTAGATGAA 1766  
|||||

QY 61 ATGTTTCCCATTAATAAGGGCGATGAGCCCTTCTCACAAAA 103  
|||||  
Db 1767 ATGTTTCCCATTAATAAGGGCGATGAGCCCTTCTCACAAAA 1809  
|||||

## RESULT 15

BF439942/c

LOCUS

DEFINITION

BF439942 641 bp mRNA linear EST 29-NOV-2000  
nacsia09.xl NCI CGAP Brn23 Homo sapiens cDNA clone IMAGE:3405976 3' similar to TR:Q9WU80 Q9WU80 CAMP INDUCIBLE 1 PROTEIN. ;, mRNA sequence.

ACCESSION

BF439942

VERSION

GI:11452459

KEYWORDS

Homo sapiens (human)

SOURCE

Homo sapiens

ORGANISM

Homo sapiens

REFERENCE

AUTHORS

TITLE

Disorders and Stroke, Brain Tumor Genome Anatomy Project

JOURNAL

COMMENT

Unpublished (1998)

Contact: Robert Strausberg, Ph.D.

Email: [cgapbs@mail.nih.gov](mailto:cgapbs@mail.nih.gov)

Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D., Ph.D.

CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL, send email to: [info@image.llnl.gov](mailto:info@image.llnl.gov)

Seq primer: -40UP from Gibco

High quality sequence stop: 450.

Location/Qualifiers

1..641

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:3405976"

/tissue\_type="glioblastoma (pooled)"

/lab\_host="DH10B"

/clone\_lib="NCI CGAP Brn23"

/note="Organ: brain; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo (dT) primer [5' TGTACCAATCTGAAGTGGAGCGCGCATATCTTTTTTTTTTTTTTTTTTTT T 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library is normalized, and was constructed by Bento Soares and M.Fatima Bonaldo."

## ORIGIN

Query Match 97.5%; Score 100.4; DB 2; Length 641;

Best Local Similarity 99.0%; Pred. No. 2e-22;

Matches 101; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGACAGCGGCGAGCTCCAGCTCTGGTTTCCTTCGTTTATCTGTTAGATGAA 60

Db 102 CGGACAGCGGCGAGCTCCAGCTCTGGTTTCCTTCGTTTATCTGTTAGATGAA 43

QY 61 ATGTTTCCCATTAATAAGGGCGATGAGCCCTTCTCACAAAA 102

Db 42 ATGTTTCCCATTAATAAGGGCGATGAGCCCTTCTCACGAAA 1

Search completed: December 7, 2005, 09:56:00

Job time : 1178.77 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 21:43:03 ; Search time 265.316 Seconds  
(without alignments)  
6427.437 Million cell updates/sec

Title: US-09-980-046B-9

Perfect score: 30

Sequence: 1 cggaaataaaggctgtgtaagacaaaaa 30

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_in.\*

3: gb\_env.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pr.\*

9: gb\_ro.\*

10: gb\_sts.\*

11: gb\_by.\*

12: gb\_un.\*

13: gb\_vl.\*

14: gb\_btg.\*

15: gb\_pl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	28.4	94.7	619	BC009507	BC009507 Homo sapi
2	28.4	94.7	668	BD203732	BD203732 Human nuc
3	28.4	94.7	668	AX014887	AX014887 Sequence
4	28.4	94.7	670	BD134423	BD134423 Human nuc
5	28.4	94.7	670	AX017252	AX017252 Sequence
6	28.4	94.7	670	AX524956	AX524956 Sequence
7	26.4	88.0	163399	AC079860	AC079860 Homo sapi
8	26.4	88.0	186759	AL645608	AL645608 Human DNA
9	26.4	88.0	200491	AB107102	AB107102 Homo sapi
10	26.4	88.0	327883	14 BX571684	BX571684 Homo sapi
11	24.4	81.3	509	CQ922970	CQ922970 Sequence
12	23.4	78.0	161151	14 AP000876	AP000876 Homo sapi
13	23.4	78.0	211945	8 AC006238	AC006238 Homo sapi
14	23	76.7	301	BD275867	BD275867 COMPOUNDS
15	23	76.7	301	AR220652	AR220652 Sequence
16	23	76.7	301	AR255646	AR255646 Sequence
17	23	76.7	301	AR281216	AR281216 Sequence
18	23	76.7	301	AR541237	AR541237 Sequence

C	19	23	76.7	301	6	AX365911	Sequence
	20	22.6	75.3	102976	15	AC142224	AC142224 Medicago
	21	22.6	75.3	199961	9	AL596111	AL596111 Mouse DNA
	22	22.2	74.0	194232	14	AC154523	AC154523 Mus muscu
C	23	22.2	74.0	208143	14	AC164357	AC164357 Bos tauru
C	24	22.2	74.0	236536	9	AC139941	AC139941 Mus muscu
	25	22	73.3	447	9	BC062266	BC062266 Mus muscu
	26	22	73.3	140864	9	AC126803	AC126803 Mus muscu
C	27	22	73.3	141855	5	CR388381	CR388381 Zebrafish
	28	22	73.3	154090	14	AC142024	AC142024 Rattus no
C	29	22	73.3	169547	9	AL590503	AL590503 Mouse DNA
	30	22	73.3	178124	9	AC153970	AC153970 Mus muscu
C	31	22	73.3	179919	14	AC114383	AC114383 Rattus no
C	32	22	73.3	181034	14	AC135486	AC135486 Rattus no
	33	22	73.3	187002	5	CR391941	CR391941 Zebrafish
C	34	22	73.3	206168	9	AC153912	AC153912 Mus muscu
	35	22	73.3	223809	14	AC105474	AC105474 Rattus no
C	36	22	73.3	238611	14	AC108953	AC108953 Rattus no
C	37	22	73.3	244255	14	AC109041	AC109041 Rattus no
C	38	22	73.3	246384	14	AC108543	AC108543 Rattus no
	39	22	73.3	265663	14	AC109741	AC109741 Rattus no
C	40	22	73.3	331363	14	AC113829	AC113829 Rattus no
C	41	21.8	72.7	189793	14	AC149188	AC149188 Papio anu
C	42	21.6	72.0	17292	8	HSEPB72EX	X85117 H. sapiens e
	43	21.6	72.0	19269	8	AL161784	AL161784 Human DNA
	44	21.6	72.0	98274	8	HS714B7	Z99755 Human DNA s
C	45	72.0	105412	14	AP007403	AP007403 Lotus cor	

## ALIGNMENTS

RESULT 1	BC009507	619 bp	mena	linear	PRI 08-MAR-2005
LOCUS	Homo sapiens interferon, alpha-inducible protein (clone IFI-15K),				
DEFINITION	mRNA (cdna clone MGC:3945 IMAGE:3545944), complete cds.				
ACCESSION	BC009507				
VERSION	BC009507.2	GI:38114701			
KEYWORDS	MGC.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 619)				
AUTHORS	Strausberg RL, Feingold EA, Grouse LH, Derge JG, Klausner RD, Collins FS, Wagner L, Shenmen CM, Schuler GD, Altschul SF, Zeeberg B, Buetow KH, Schaefer CF, Bhat NK, Hopkins RF, Jordan H, Moore T, Max SI, Wang J, Hsieh F, Diatchenko L, Marusina K, Farmer AA, Rubin GM, Hong L, Stapleton M, Soares MB, Bonaldo MF, Casavant TL, Scheetz TE, Brownstein MJ, Usdin TB, Toshiyuki S, Carninci P, Prange C, Raha SS, Loquellano NA, Peters GJ, Abramson RD, Mullaly SJ, Bosak SA, McEwan PJ, McKernan KJ, Malek JA, Gunaratne PH, Richards S, Worley KC, Hale S, Garcia AM, Gay LJ, Hulyk SW, Villalón DK, Muzny DM, Sodergren EJ, Lu X, Gibbs RA, Fahey J, Helton E, Kertanen M, Madan A, Rodriguez S, Sanchez A, Whiting M, Madan A, Young AC, Shevchenko Y, Bouffard GG, Blakeley RM, Touchman JW, Green ED, Dickson MC, Rodriguez AC, Grimwood J, Schmutz J, Myers RM, Butterfield JS, Krzywinski MI, Skalska U, Smalusz DE, Schnerch A, Schein JF, Jones SJ and Marra MA.				
CONSRTM	Mammalian Gene Collection Program Team				
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)				
PUBMED	12477932				
REFERENCE	2 (bases 1 to 619)				
AUTHORS	Director MGC project.				
TITLE	Direct Submission				
JOURNAL	Submitted (25-JUN-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA				

REMARK N1H-MGC Project URL: <http://mgc.nci.nih.gov>  
 COMMENT On Oct 31, 2003 this sequence version replaced gi:14550513.  
 Contact: MGC help desk  
 Email: [cgapbs-rc@mail.nih.gov](mailto:cgapbs-rc@mail.nih.gov)  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Institute for Systems Biology  
<http://www.systemsbio.org>  
 contact: [amadansystemsbiology.org](mailto:amadansystemsbiology.org)  
 Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAL Plate: 11 Row: b Column: 8  
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4826773.

FEATURES  
 source Location/Qualifiers  
 1..619  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="MGC:3945 IMAGE:3545944"  
 /tissue\_type="Colon, adenocarcinoma"  
 /clone\_1lb="NIH\_MGC\_15"  
 /lab\_host="DH10B-R"  
 /note="Vector: pOTB7"  
 1..619  
 /gene="G1P2"  
 /note="synonyms: ISG15, IFI15, UCRP"  
 /db\_xref="GeneID:9636"  
 /db\_xref="MIM:147571"  
 42..539  
 /gene="G1P2"  
 /codon\_start=1  
 /product="interferon, alpha-inducible protein (clone IFI-15K)"  
 /protein\_id="AAH09507.1"  
 /db\_xref="GI:14550514"  
 /db\_xref="GeneID:9636"  
 /db\_xref="MIM:147571"  
 /translation="MGDLTVKMLAGNEFQVLSSSMSVSELKQITKIGVHAFQOR LAVHPSVALQDRVPLASQGLPGSTVLLVVDKDEPLNILVRNKKRSTYVRUTQ TVAHLKQGVSGLEGVQDDLFWLTFEGKPLEDQLPGYGLKPLSTVPMNLRRLRGGTE PGGRS"

ORIGIN  
 Query Match 94.7%; Score 28.4; DB 8; Length 619;  
 Best Local Similarity 96.7%; Pred. No. 5.1; Mismatches 0; Indels 0; Gaps 0;  
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 1 CGGAATAAAGCGCTGTTGTTAAAGACAAAA 30  
 |||||  
 574 CGGAATAAAGCGCTGTTGTTAAAGACAAAA 603

RESULT 2  
 BD203732  
 LOCUS 668 bp DNA linear PAT 17-JUL-2003  
 DEFINITION Human nucleic acid sequence originating in ovarian myoma tissue.  
 ACCESSION BD203732  
 VERSION BD203732.1 GI:33013502  
 KEYWORDS JP 2002511252-A/73.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 668)  
 Specht,T., Hinzmann,B., Schmitt,A., Pilarsky,C., Dahl,E. and Rosenthal,A.  
 Human nucleic acid sequence originating in ovarian myoma tissue

REFERENCE  
 AUTHORS  
 TITLE

JOURNAL N1H-MGC Project URL: <http://mgc.nci.nih.gov>  
 COMMENT On Oct 31, 2003 this sequence version replaced gi:14550513.  
 Contact: MGC help desk  
 Email: [cgapbs-rc@mail.nih.gov](mailto:cgapbs-rc@mail.nih.gov)  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Institute for Systems Biology  
<http://www.systemsbio.org>  
 contact: [amadansystemsbiology.org](mailto:amadansystemsbiology.org)  
 Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAL Plate: 11 Row: b Column: 8  
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4826773.

FEATURES  
 source Location/Qualifiers  
 1..619  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="MGC:3945 IMAGE:3545944"  
 /tissue\_type="Colon, adenocarcinoma"  
 /clone\_1lb="NIH\_MGC\_15"  
 /lab\_host="DH10B-R"  
 /note="Vector: pOTB7"  
 1..619  
 /gene="G1P2"  
 /note="synonyms: ISG15, IFI15, UCRP"  
 /db\_xref="GeneID:9636"  
 /db\_xref="MIM:147571"  
 42..539  
 /gene="G1P2"  
 /codon\_start=1  
 /product="interferon, alpha-inducible protein (clone IFI-15K)"  
 /protein\_id="AAH09507.1"  
 /db\_xref="GI:14550514"  
 /db\_xref="GeneID:9636"  
 /db\_xref="MIM:147571"  
 /translation="MGDLTVKMLAGNEFQVLSSSMSVSELKQITKIGVHAFQOR LAVHPSVALQDRVPLASQGLPGSTVLLVVDKDEPLNILVRNKKRSTYVRUTQ TVAHLKQGVSGLEGVQDDLFWLTFEGKPLEDQLPGYGLKPLSTVPMNLRRLRGGTE PGGRS"

ORIGIN  
 Query Match 94.7%; Score 28.4; DB 8; Length 619;  
 Best Local Similarity 96.7%; Pred. No. 5.1; Mismatches 0; Indels 0; Gaps 0;  
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 1 CGGAATAAAGCGCTGTTGTTAAAGACAAAA 30  
 |||||  
 574 CGGAATAAAGCGCTGTTGTTAAAGACAAAA 603

RESULT 2  
 BD203732  
 LOCUS 668 bp DNA linear PAT 17-JUL-2003  
 DEFINITION Human nucleic acid sequence originating in ovarian myoma tissue.  
 ACCESSION BD203732  
 VERSION BD203732.1 GI:33013502  
 KEYWORDS JP 2002511252-A/73.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 668)  
 Specht,T., Hinzmann,B., Schmitt,A., Pilarsky,C., Dahl,E. and Rosenthal,A.  
 Human nucleic acid sequence originating in ovarian myoma tissue

REFERENCE  
 AUTHORS  
 TITLE

JOURNAL N1H-MGC Project URL: <http://mgc.nci.nih.gov>  
 COMMENT On Oct 31, 2003 this sequence version replaced gi:14550513.  
 Contact: MGC help desk  
 Email: [cgapbs-rc@mail.nih.gov](mailto:cgapbs-rc@mail.nih.gov)  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Institute for Systems Biology  
<http://www.systemsbio.org>  
 contact: [amadansystemsbiology.org](mailto:amadansystemsbiology.org)  
 Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAL Plate: 11 Row: b Column: 8  
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4826773.

FEATURES  
 source Location/Qualifiers  
 1..668  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 1 CGGAATAAAGCGCTGTTGTTAAAGACAAAA 30  
 |||||  
 637 CGGAATAAAGCGCTGTTGTTAAAGACAAAA 666

ORIGIN  
 Query Match 94.7%; Score 28.4; DB 6; Length 668;  
 Best Local Similarity 96.7%; Pred. No. 5; Mismatches 0; Indels 0; Gaps 0;  
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 1 CGGAATAAAGCGCTGTTGTTAAAGACAAAA 30  
 |||||  
 637 CGGAATAAAGCGCTGTTGTTAAAGACAAAA 666

RESULT 3  
 AX014887  
 LOCUS 668 bp DNA linear PAT 07-SEP-2000  
 DEFINITION Sequence 85 from Patent WO9953040.  
 ACCESSION AX014887  
 VERSION AX014887.1 GI:10041154  
 KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.  
 1  
 Schmitt,A., Specht,T., Dahl,E., Hinzmann,B., Rosenthal,A. and Pilarsky,C.  
 Human nucleic acid sequences from ovarian tumour tissue  
 Patent: WO 9953040-A 85 21-OCT-1999;  
 Schmitt ARMIN (DE); SPECHT THOMAS (DE); DAHL EDGAR (DE); HINZMANN BERND (DE); ROSENTHAL ANDRE (DE); METAGEN GES FUER GENOMFORSCHUN (DE); PILARSKY CHRISTIAN (DE)  
 Location/Qualifiers  
 1..668  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"

ORIGIN  
 Query Match 94.7%; Score 28.4; DB 6; Length 668;  
 Best Local Similarity 96.7%; Pred. No. 5; Mismatches 0; Indels 0; Gaps 0;  
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 1 CGGAATAAAGCGCTGTTGTTAAAGACAAAA 30  
 |||||  
 637 CGGAATAAAGCGCTGTTGTTAAAGACAAAA 666

RESULT 4

```

BD134423
LOCUS       BD134423               670 bp      DNA      linear      PAT 18-SEP-2002
DEFINITION   Human nucleic acid sequence originating in mammary tumor tissue.
ACCESSION    BD134423
VERSION      BD134423.1  GI:23229368
KEYWORDS     JP 2002506643-A/1.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
             Hominiidae; Homo.
REFERENCE    1 (bases 1 to 670)
AUTHORS     Speft,T., Hintzman,B., Armin,S., Pirarski,C., Edgar,D. and
             Rosenthal,A.
TITLE       Human nucleic acid sequence originating in mammary tumor tissue
JOURNAL     Patent: JP 2002506643-A 1 05-MAR-2002;
            METAGEN GESELLSCHAFT FUER GENOME FORSCHUNG MBH
COMMENT     OS Homo sapiens (human)
            PN JP 2002506643-A/1
            PD 05-MAR-2002
            PF 19-MAR-1999 JP 2000536852
            PR 20-MAR-1998 DE 198 13 839.3
            PI THOMAS SPEFT, BERND HINTZMAN, SHCMITT ARMIN, CHRISTIAN PIRARSKI,
            DUHL EDGAR,
            PI ANDRE ROSENTHAL
            PC C12N15/09,A61K31/711,A61K38/00,A61K48/00,A61P35/00,C07K14/47,
            C07K16/18,
            PC C12N15/10,C12N15/00,A61K37/02,C12N5/00 CC Human
            nucleic acid sequence originating in mammary tumor CC
            tissue
FEATURES     source
             FT Key Location/Qualifiers
             FT source 1..670
             /organism="Homo sapiens (human)".
             /mol_type="genomic DNA"
             /db_xref="taxon:9606"

ORIGIN
Query Match 94.7%; Score 28.4; DB 6; Length 670;
Best Local Similarity 96.7%; Pred. No. 5;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 CGGAATAAAGCGCTGTTGTAAGACAAAAA 30
    |||||
Db 637 CGGAATAAAGCGCTGTTGTAAGACAAAAA 666

RESULT 5
AC017252
LOCUS       AX017252               670 bp      DNA      linear      PAT 07-SEP-2000
DEFINITION   Sequence 2 from Patent WO9947669.
ACCESSION    AX017252
VERSION      AX017252.1  GI:10042170
KEYWORDS     .
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
             Hominiidae; Homo.
REFERENCE    1
AUTHORS     Schmitt,A., Specht,T., Dahl,E., Hinzmann,B., Rosenthal,A. and
             Pilarsky,C.
TITLE       Human nucleic acid sequences from tissue of breast tumors
JOURNAL     Patent: WO 9947669-A 2 23-SEP-1999;
            SCHMITT ARMIN (DE); SPECHT THOMAS (DE); DAHL EDGAR (DE); HINZMANN
            BERND (DE); ROSENTHAL ANDRE (DE); METAGEN GES FUER GENOMFORSCHUN
            (DE); PILARSKY CHRISTIAN (DE)
FEATURES     source
             1..670
             /organism="Homo sapiens"
             /mol_type="unassigned DNA"

BD134423
LOCUS       BD134423               670 bp      DNA      linear      PAT 18-SEP-2002
DEFINITION   Human nucleic acid sequence originating in mammary tumor tissue.
ACCESSION    BD134423
VERSION      BD134423.1  GI:23229368
KEYWORDS     JP 2002506643-A/1.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
             Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
             Hominiidae; Homo.
REFERENCE    1 (bases 1 to 670)
AUTHORS     Speft,T., Hintzman,B., Armin,S., Pirarski,C., Edgar,D. and
             Rosenthal,A.
TITLE       Human nucleic acid sequence originating in mammary tumor tissue
JOURNAL     Patent: JP 2002506643-A 1 05-MAR-2002;
            METAGEN GESELLSCHAFT FUER GENOME FORSCHUNG MBH
COMMENT     OS Homo sapiens (human)
            PN JP 2002506643-A/1
            PD 05-MAR-2002
            PF 19-MAR-1999 JP 2000536852
            PR 20-MAR-1998 DE 198 13 839.3
            PI THOMAS SPEFT, BERND HINTZMAN, SHCMITT ARMIN, CHRISTIAN PIRARSKI,
            DUHL EDGAR,
            PI ANDRE ROSENTHAL
            PC C12N15/09,A61K31/711,A61K38/00,A61K48/00,A61P35/00,C07K14/47,
            C07K16/18,
            PC C12N15/10,C12N15/00,A61K37/02,C12N5/00 CC Human
            nucleic acid sequence originating in mammary tumor CC
            tissue
FEATURES     source
             FT Key Location/Qualifiers
             FT source 1..670
             /organism="Homo sapiens"
             /mol_type="genomic DNA"
             /db_xref="taxon:9606"

ORIGIN
Query Match 94.7%; Score 28.4; DB 6; Length 670;
Best Local Similarity 96.7%; Pred. No. 5;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 CGGAATAAAGCGCTGTTGTAAGACAAAAA 30
    |||||
Db 637 CGGAATAAAGCGCTGTTGTAAGACAAAAA 666

RESULT 7
AC079860/c
LOCUS       AC079860               163399 bp     DNA      linear      HTG 23-DEC-2000
DEFINITION   Homo sapiens chromosome RPCI-11 clone RP11-671C15, WORKING DRAFT
            SEQUENCE, 54 unordered pieces.
ACCESSION    AC079860
VERSION      AC079860.3  GI:11990750
KEYWORDS     HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
             Hominiidae; Homo.
REFERENCE    1 (bases 1 to 163399)
AUTHORS     Waterston,R.H.
TITLE       The sequence of Homo sapiens clone
JOURNAL     Unpublished
REFERENCE    2 (bases 1 to 163399)
AUTHORS     Waterston,R.H.
TITLE       Direct Submission
JOURNAL     Submitted (14-SEP-2000) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On Dec 23, 2000 this sequence version replaced gi:10518428.
            ----- Genome Center -----
            Center: Washington University Genome Sequencing Center

```

```

ORIGIN
Query Match 94.7%; Score 28.4; DB 6; Length 670;
Best Local Similarity 96.7%; Pred. No. 5;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 CGGAATAAAGCGCTGTTGTAAGACAAAAA 30
    |||||
Db 637 CGGAATAAAGCGCTGTTGTAAGACAAAAA 666

RESULT 6
AX524956
LOCUS       AX524956               670 bp      DNA      linear      PAT 21-NOV-2002
DEFINITION   Sequence 2 from Patent EPI236799.
ACCESSION    AX524956
VERSION      AX524956.1  GI:25170038
KEYWORDS     .
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
             Hominiidae; Homo.
REFERENCE    1
AUTHORS     Specht,T., Hinzmann,B., Schmitt,A., Pilarsky,C., Dahl,E. and
             Rosenthal,A.
TITLE       Human nucleic acid sequences derived from breast tumor tissue
JOURNAL     Patent: EP 1236799-A 2 04-SEP-2002;
            metagen Pharmaceuticals GmbH (DE)
FEATURES     source
             1..670
             /organism="Homo sapiens"
             /mol_type="unassigned DNA"
             /db_xref="taxon:9606"

ORIGIN
Query Match 94.7%; Score 28.4; DB 6; Length 670;
Best Local Similarity 96.7%; Pred. No. 5;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 CGGAATAAAGCGCTGTTGTAAGACAAAAA 30
    |||||
Db 637 CGGAATAAAGCGCTGTTGTAAGACAAAAA 666

RESULT 7
AC079860/c
LOCUS       AC079860               163399 bp     DNA      linear      HTG 23-DEC-2000
DEFINITION   Homo sapiens chromosome RPCI-11 clone RP11-671C15, WORKING DRAFT
            SEQUENCE, 54 unordered pieces.
ACCESSION    AC079860
VERSION      AC079860.3  GI:11990750
KEYWORDS     HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
             Hominiidae; Homo.
REFERENCE    1 (bases 1 to 163399)
AUTHORS     Waterston,R.H.
TITLE       The sequence of Homo sapiens clone
JOURNAL     Unpublished
REFERENCE    2 (bases 1 to 163399)
AUTHORS     Waterston,R.H.
TITLE       Direct Submission
JOURNAL     Submitted (14-SEP-2000) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On Dec 23, 2000 this sequence version replaced gi:10518428.
            ----- Genome Center -----
            Center: Washington University Genome Sequencing Center

```

Center code: WUGSC  
 Web site: http://genome.wustl.edu/gsc/index.shtml  
 ----- Project Information -----  
 Center project name: H1NH0671C15  
 ----- Summary Statistics -----  
 Sequencing vector: M13; 100%  
 Sequencing vector: plasmid; 0%  
 Chemistry: Dye-terminator Big Dye; 0% of reads  
 Chemistry: Dye-terminator Big Dye; 0% of reads  
 Assembly program: Phrap; version 0.990319  
 Consensus quality: 123674 bases at least Q40  
 Consensus quality: 137441 bases at least Q30  
 Consensus quality: 143762 bases at least Q20  
 Insert size: 176000; agarose-fp  
 Insert size: 158099; sum-of-contigs  
 Quality coverage: 2.94 in Q20 bases; agarose-fp  
 Quality coverage: 3.43 in Q20 bases; sum-of-contigs  
 -----  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 54 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

1 1229: contig of 1229 bp in length  
 1230: gap of unknown length  
 1231: contig of 1115 bp in length  
 1232: gap of unknown length  
 1233: contig of 1637 bp in length  
 1234: gap of unknown length  
 1235: contig of 1297 bp in length  
 1236: gap of unknown length  
 1237: contig of 1052 bp in length  
 1238: gap of unknown length  
 1239: contig of 1306 bp in length  
 1240: gap of unknown length  
 1241: contig of 1303 bp in length  
 1242: gap of unknown length  
 1243: contig of 1606 bp in length  
 1244: gap of unknown length  
 1245: contig of 1403 bp in length  
 1246: gap of unknown length  
 1247: contig of 1546 bp in length  
 1248: gap of unknown length  
 1249: contig of 1401 bp in length  
 1250: gap of unknown length  
 1251: contig of 1088 bp in length  
 1252: gap of unknown length  
 1253: contig of 1230 bp in length  
 1254: gap of unknown length  
 1255: contig of 1979 bp in length  
 1256: gap of unknown length  
 1257: contig of 1586 bp in length  
 1258: gap of unknown length  
 1259: contig of 1851 bp in length  
 1260: gap of unknown length  
 1261: contig of 2326 bp in length  
 1262: gap of unknown length  
 1263: contig of 2345 bp in length  
 1264: gap of unknown length  
 1265: contig of 1900 bp in length  
 1266: gap of unknown length  
 1267: contig of 1578 bp in length  
 1268: gap of unknown length  
 1269: contig of 1435 bp in length  
 1270: gap of unknown length  
 1271: contig of 3205 bp in length  
 1272: gap of unknown length  
 1273: contig of 2565 bp in length  
 1274: gap of unknown length  
 1275: contig of 1545 bp in length

41829: gap of unknown length  
 41830: contig of 3035 bp in length  
 41831: gap of unknown length  
 41832: contig of 3279 bp in length  
 41833: gap of unknown length  
 41834: contig of 1666 bp in length  
 41835: gap of unknown length  
 41836: contig of 2322 bp in length  
 41837: gap of unknown length  
 41838: contig of 2331 bp in length  
 41839: gap of unknown length  
 41840: contig of 2490 bp in length  
 41841: gap of unknown length  
 41842: contig of 2256 bp in length  
 41843: gap of unknown length  
 41844: contig of 2111 bp in length  
 41845: gap of unknown length  
 41846: contig of 4216 bp in length  
 41847: gap of unknown length  
 41848: contig of 3361 bp in length  
 41849: gap of unknown length  
 41850: contig of 3354 bp in length  
 41851: gap of unknown length  
 41852: contig of 2816 bp in length  
 41853: gap of unknown length  
 41854: contig of 3328 bp in length  
 41855: gap of unknown length  
 41856: contig of 2931 bp in length  
 41857: gap of unknown length  
 41858: contig of 2884 bp in length  
 41859: gap of unknown length  
 41860: contig of 3816 bp in length  
 41861: gap of unknown length  
 41862: contig of 4009 bp in length  
 41863: gap of unknown length  
 41864: contig of 3323 bp in length  
 41865: gap of unknown length  
 41866: contig of 3662 bp in length  
 41867: gap of unknown length  
 41868: contig of 3956 bp in length  
 41869: gap of unknown length  
 41870: contig of 5748 bp in length  
 41871: gap of unknown length  
 41872: contig of 4714 bp in length  
 41873: gap of unknown length  
 41874: contig of 4740 bp in length  
 41875: gap of unknown length  
 41876: contig of 5098 bp in length  
 41877: gap of unknown length  
 41878: contig of 6020 bp in length  
 41879: gap of unknown length  
 41880: contig of 6224 bp in length  
 41881: gap of unknown length  
 41882: contig of 5412 bp in length  
 41883: gap of unknown length  
 41884: contig of 5024 bp in length  
 41885: gap of unknown length  
 41886: contig of 7245 bp in length  
 41887: gap of unknown length  
 41888: contig of 7200 bp in length.

## FEATURES

source

Location/Qualifiers

1..163399  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosomes="RPC1-11"  
 /clone="Rp11-671C15"  
 1..1229  
 /notes="assembly\_name:Contig22"  
 1230..1329  
 /estimated\_length=unknown  
 1330..2444  
 /notes="assembly\_name:Contig25"

misc\_feature

gap

misc\_feature

```

gap          2445..2544
/estimated_length=unknown
misc_feature 2545..4181
/notes="assembly_name:Contig29"
gap          4182..4281
/estimated_length=unknown
misc_feature 4282..5578
/notes="assembly_name:Contig30"
gap          5579..5678
/estimated_length=unknown
misc_feature 5679..6730
/notes="assembly_name:Contig31"
gap          6731..6830
/estimated_length=unknown
misc_feature 6831..8136
/notes="assembly_name:Contig32"
gap          8137..8236
/estimated_length=unknown
misc_feature 8237..9539
/notes="assembly_name:Contig33"
gap          9540..9639
/estimated_length=unknown
misc_feature 9640..11245
/notes="assembly_name:Contig34"
gap          11246..11345
/estimated_length=unknown
misc_feature 11346..12748
/notes="assembly_name:Contig35"
gap          12749..12848

```

Query Match 88.0%; Score 26.4; DB 14; Length 163399;

Best Local Similarity 96.4%; Pred. No. 8;

Matches 27; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGAATAAAGCGTGTGTAAGACAAA 28

Db 120846 CGGAATAAAGCGTGTGTAAGACAAA 120819

# RESULT 8

LOCUS

DEFINITION AL645608 186759 bp DNA linear PRI 18-MAY-2005  
Human DNA sequence from clone RP11-5407 on chromosome 1. Contains  
three novel genes, a novel gene (MGC45873), a novel gene  
(DKFZP564C186), a novel gene, a novel gene (DKFZP434H2010), a novel  
gene (MGC13275), the gene for bHLH factor Hes4 (Hes4), a ribosomal  
protein L39 (RPL39) pseudogene, a novel gene, the GIP2 gene for  
interferon alpha-inducible protein (clone IFI-15K), the AGRN gene  
for agrin, two novel genes and fifteen CpG islands, complete  
sequence.

ACCESSION AL645608

VERSION AL645608.30 GI:48374144

KEYWORDS HTG.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 186759)

REFERENCE

AUTHORS

TITLE Direct Submission

JOURNAL

Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk

Clone requests: clonerequest@sanger.ac.uk

On Jun 6, 2004 this sequence version replaced gi:45125204.

The following abbreviations are used to associate primary accession

numbers given in the feature table with their source databases:

Em, EMBL; Sw, SWISSPROT; Tr, TrEMBL; Wp, WORMPEP; Information

on the WORMPEP database can be found at

http://www.sanger.ac.uk/projects/c\_elegans/wormpep This sequence

was generated from part of bacterial clone contigs of human

chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping

Group. Further information can be found at

http://www.sanger.ac.uk/HGP/Chr1  
RP11-5407 is from the library RP11-11.1 constructed by the group of  
Pietér de Jong. For further details see  
http://www.chori.org/bacpac/home.htm  
VECTOR: pBACE3.6

----- Genome Center

Center: Wellcome Trust Sanger Institute

Center code: SC

Web site: http://www.sanger.ac.uk

Contact: vegas@sanger.ac.uk

-----

This sequence was finished as follows unless otherwise noted: all  
regions were either double-stranded or sequenced with an alternate  
chemistry or covered by high quality data (i.e., phred quality >=  
30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by at least  
one subclone; and the assembly was confirmed by restriction digest,  
except on the rare occasion of the clone being a YAC.

## FEATURES

### Source

```

1..186759
  Location/Qualifiers
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    /chromosome="1"
    /clone="RP11-5407"
    /clone_lib="RP11-11.1"
954
  misc_feature
    /note="Clone right end: RP11-206L10"
    join(<955..956,63125..63310,64994..65156,65986..66101,
66260..66338,66409..66908,67103..67227,67548..68109)
    /locus_tag="RP11-5407.4-004"
    join(<955..956,63125..63310,64994..65156,65986..66101,
66260..66338,66409..66908,67103..67227,67548..68109)
    /locus_tag="RP11-5407.4-004"
    /product="novel protein"
    /note="match: CDNAS: Em:AK054643.1"
    join(<955..956,63125..63310,64994..65156,65986..66101,
66260..66338,66409..66908,67103..67227,67548..68003)
    /locus_tag="RP11-5407.4-004"
    /standard_name="OTTHUMP0000038430"
    /note="match: proteins: Q96NU1"
    /codon_start=2
    /product="novel protein"
    /protein_id="CAI15563.1"
    /db_xref="GI:55962171"
    /db_xref="InterPro:IPR001560"
    /db_xref="InterPro:IPR011510"
    /db_xref="UniProt/TREMBL:O5SVA0"
    /translation="EALLPRELGSNAPEDHYRLVLSALSEASTFEDPQRLYHLGLP
SHGEDPPHDPHPLPSHDLRVQEVAAALRGPSGLEAHLPSSTAGQRKQGLAQH
REGAAPAAAPFSERELPQPPILLSPONAPHVALGPHLRPFLGVPALCQTPGQFL
PPQAEMFANQQLLRKQNLARLELPADLLRQKELASARQLLAPETALRPNDGAEEL
ORRGALLVNHGAAPLLALPPQPPGSPPTPSRDSARRAPKGGPGPGAPARPSSEKE
MTGARLWAQDSEDEPPKSDGEDPTAAVCGRGPTPGQAPAGAGAGKGLFPFSGTL
PLGFPYAVSPYFHTGVLGSMGDEAPEDVTWKTVDDVCSFVGLSGCGEYTRVP
REGIDGETLPLTEHLLTMGLKLPALKIAQVRRWGVRSQPHSVAESGWCVC
DSHQALSLQVARRLGRVFVYASFPVALPQPPTLRAPERLGTGEQPLSPPTATSPY
DGGHALAGQTSKQENGTALLPGAPDPSQPLC"
    join(<30099..30107,59622..59746,62890..62979,63125..63262,
64994..65156,65986..66101,66260..66338,66409..66908,
67103..67227,67548..67658,67758..68424)
    /locus_tag="RP11-5407.4-001"
    join(<30039..30107,59622..59746,62890..62979,63125..63262,
64994..65156,65986..66101,66260..66338,66409..66908,
67103..67227,67548..67658,67758..68424)
    /locus_tag="RP11-5407.4-001"
    /product="novel protein"
    /note="match: ESTs: Em:BE258987.1 Em:CA389330.1
match: CDNAS: Em:BC024295.2"
    join(<30099..30107,59622..59746,62890..62979,63125..63262,
64994..65156,65986..66101,66260..66338,66409..66908,
67103..67227,67548..67658,67758..68003)
    /locus_tag="RP11-5407.4-001"

```

```
/standard_name="OTTHUMP00000038427"
/notes="match: proteins: Tr:Q8TB59"
/codon_start=2
/product="novel protein"
/protein_id="CAI15564.1"
/db_xref="GI:55962172"
/db_xref="InterPro:IPR001660"
/db_xref="UniProt:IPR011510"
/db_xref="UniProt/Tx:EMBL:O5SV99"
/translation="SCSDGNLPTLTISSVRSRLVMPHEQSRCEQRGSLEIGLRPA
GDLKRLGRSPRISSDCFKRARSPOEALLPRELPGSPMAPEDHYRRLVSAUSE
ASTFEDQRHYHLGSHDLRLVRQEVAAALRPGSLGLEAHLPSSTAGQRKQGLAQH
REGAFAAALPFPPLPSQNAAPHVALPGHLRPFPGVFSALCQTPGYGFL
PQAQAEFAWQELLRKQNLARLEPADLLKQELSRAPOLLAPETALRPNDSAEEL
ORGLALVNHGAAPLALPQGPSPGPPPTPSRDSARRAPKGGPGPASARSESKE
MTGNLWAQGSDEPKDSGDGDPETAACVCGRPTFGQAPAGAGAEKGLFPQSTL
PLGPPYAPFYFHFGAVGLSMDGEEAPEDVTKWTDVDCFVGLSGCGGBYTRVF
REGQIDETLPLTEEHLHLMGLKLPALKIRQAVRRGRVYVASFVAPLQPP
TLRAPRELGTGQPLSPPTTATSPYGGGHALAGTSPKQENGTLALLPGAPDPSQLC
"
32390. .32684
/locus_tag="RP11-5407.4-001"
/notes="single clone region. Sequence generated from a
short insert library derived from a single pUC clone.
Restriction digest data confirm the assembly."
join(35285. .35323,35795. .38798)
/locus_tag="RP11-5407.1-001"
join(35285. .35323,35795. .38798)
/locus_tag="RP11-5407.1-001"
/product="novel transcript"
/notes="match: cDNAs: Em:AK056486.1"
join(38021. .38294,38653. .38821)
/locus_tag="RP11-5407.2-001"
join(38021. .38294,38653. .38821)
/locus_tag="RP11-5407.2-001"
/product="putative novel transcript"
/notes="match: ESTs: Em:AA633998.1"
complement(join(40720. .41570,41872. .41960,42675. .42765,
43185. .43542))
/locus_tag="RP11-5407.3-001"
complement(join(40720. .41570,41872. .41960,42675. .42765,
43185. .43542))
/locus_tag="novel transcript"
/notes="match: ESTs: Em:BU619380.1 Em:BU632785.1
Em:BU738526.1 Em:CA447406.1"
complement(40720)
/locus_tag="RP11-5407.3-001"
complement(40741. .40746)
/locus_tag="RP11-5407.3-001"
complement(join(41220. .41336,41872. .42384))
/locus_tag="RP11-5407.3-002"
complement(join(41220. .41336,41872. .42384))
/locus_tag="novel transcript"
/notes="match: ESTs: Em:BM758684.1 Em:BM759761.1
Em:BM765468.1"
complement(41423)
/locus_tag="RP11-5407.3-002"
complement(41440. .41445)
/locus_tag="RP11-5407.3-002"
join(<49000. .49039,49772. .49863,54005. .54186,54889. .54939,
59622. .>59643)
/locus_tag="RP11-5407.4-003"
join(<49000. .49039,49772. .49863,54005. .54186,54889. .54939,
59622. .>59643)
/product="novel protein"
/notes="match: ESTs: Em:AA427964.1"
join(<49000. .49039,49772. .49863,54005. .54186,54889. .54939,
59622. .>59643)
/locus_tag="RP11-5407.4-003"
/standard_name="OTTHUMP00000038429"

miac_feature
/standard_name="OTTHUMP00000038427"
/notes="match: proteins: Tr:Q8TB59"
/codon_start=2
/product="novel protein"
/protein_id="CAI15565.1"
/db_xref="GI:55962173"
/db_xref="UniProt/Tx:EMBL:O5SV95"
/translation="ORLGLAGHSLRSEKSLKTLMSKGLQVHPPICDPCGCRISPV
NRGRGLAKRTVALPAARNLKRTSPFSASDGDSDGSGPTCGRRRGLKQDGPHIRIM
KRRVTHWDVNISFREASCSQGNLPT"
join(<49601. .49650,49772. .49863,54005. .54186,54889. .54939,
59622. .59746,62890. .>62972)
/locus_tag="RP11-5407.4-002"
join(<49601. .49650,49772. .49863,54005. .54186,54889. .54939,
59622. .59746,62890. .>62972)
/locus_tag="RP11-5407.4-002"
/product="novel protein"
/notes="match: ESTs: Em:CA396407.1"
join(<49601. .49650,49772. .49863,54005. .54186,54889. .54939,
59622. .59746,62890. .>62972)
/locus_tag="RP11-5407.4-002"
/standard_name="OTTHUMP00000038428"
/codon_start=2
/product="novel protein"

Query Match 88.0%; Score 26.4; DB 8; Length 186759;
Best Local Similarity 96.4%; Pred. No. 7.8;
Matches 27; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGAATAAAGGCTGTTGTAAGACAAA 28
|||||
DB 138363 CGGAATAAAGGCTGTTGTAAGAGAAA 138390

RESULT 9
AB107102
LOCUS
DEFINITION
Homo sapiens chromosome 1 clone RP11-671C15 map lp36.3, ***
SEQUENCING IN PROGRESS ***, 13 unordered pieces.
AB107102
VERSION
AB107102.1 GI:33636391
HTG; HTGS_PHASE1.
KEYWORDS
Homo sapiens
SOURCE
Homo sapiens
ORGANISM
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1
AUTHORS
Hosoya,N., Ogawa,S., Motokura,T., Hangaishi,A., Wang,L., Qiao,Y.,
Nannya,Y., Kogi,M. and Hirai,H.
TITLE
Molecular cytogenetic analyses of HIG, a novel human cell line
carrying t(1;3)(p36.3;q25.3) established from a patient with
chronic myelogenous leukemia in blastic crisis
JOURNAL
Int. J. Hematol. 78 (5), 432-438 (2003)
PUBMED
14704036
REFERENCE
2 (bases 1 to 200491)
Hosoya,N.
Direct Submission
TITLE
Submitted (28-MAR-2003) Noriko Hosoya, University of Tokyo.
JOURNAL
Department of Hematology and Oncology; Hongo 7-3-1, Bunkyo-ku,
Tokyo 113-8655, Japan (E-mail:hosoya-tyk@umin.ac.jp,
Tel:81-3-3815-5411(ex.35602), Fax:81-3-5804-6261)
COMMENT
* NOTE: This is a 'working draft' sequence. It currently
* consists of 13 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.
* 1 35799: contig of 35799 bp in length
* 35800 35899: gap of 100 bp
* 35900 39521: contig of 3622 bp in length
* 39522 39621: gap of 100 bp
* 39622 53971: contig of 14350 bp in length
```



\* 53972 54071: gap of 100 bp  
 \* 54072 89894: contig of 35823 bp in length  
 \* 89895 89994: gap of 100 bp  
 \* 89995 109993: contig of 19999 bp in length  
 \* 109994 110093: gap of 100 bp  
 \* 110094 124472: contig of 14379 bp in length  
 \* 124473 124572: gap of 100 bp  
 \* 124573 128857: contig of 4285 bp in length  
 \* 128858 128957: gap of 100 bp  
 \* 128958 139050: contig of 10093 bp in length  
 \* 139051 139150: gap of 100 bp  
 \* 139151 145733: contig of 6583 bp in length  
 \* 145734 145833: gap of 100 bp  
 \* 145834 156433: contig of 10600 bp in length  
 \* 156434 156533: gap of 100 bp  
 \* 156534 177511: contig of 20978 bp in length  
 \* 177512 177611: gap of 100 bp  
 \* 177612 186275: contig of 8664 bp in length  
 \* 186276 186375: gap of 100 bp  
 \* 186376 200491: contig of 14116 bp in length.

## FEATURES

source

1..200491  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="1"  
 /map="1p36.3"  
 /clone="RP11-671C15"

## ORIGIN

Query Match 88.0%; Score 26.4; DB 14; Length 200491;  
 Best Local Similarity 96.4%; Pred. No. 7.7;  
 Matches 27; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGAATAAAGCGTGTGTTAAAGACAAA 28

Db 150670 CGGAATAAAGCGTGTGTTAAAGACAAA 150697

## RESULT 10

LOCUS BX571684 327883 bp DNA linear HTG 15-JUL-2003  
 DEFINITION Homo sapiens chromosome 1 clone CTD-3032124, 3 unordered pieces.  
 ACCESSION BX571684  
 VERSION BX571684.2 GI:32812682  
 KEYWORDS HTG; HTGS\_PHASE1; HTGS\_CANCELLED.  
 SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homidae; Homo

## REFERENCE

1 (bases 1 to 327883)  
 Hall, R.  
 Direct Submission  
 Submitted (14-JUL-2003) Wellcome Trust Sanger Institute, Hinxton,  
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
 humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk

## COMMENT

On Jul 15, 2003 this sequence version replaced gi:32567439.  
 ----- Genome Center  
 Center: Wellcome Trust Sanger Institute  
 Center code: SC

Web site: <http://www.sanger.ac.uk>  
 Contact: humquery@sanger.ac.uk  
 ----- Project Information  
 Center project name: BX3032124  
 ----- Summary Statistics  
 Assembly program: XGAP4; version 4.5  
 Chemistry: Dye-terminator; 100% of reads  
 Consensus quality: 184256 bases at least Q40  
 Consensus quality: 184906 bases at least Q30  
 Consensus quality: 185360 bases at least Q20  
 Insert size: 327683; sum-of-contigs  
 Insert size: 195389; 4.5% error; agarose-fp

Quality coverage: 4.02x in Q20 bases; sum-of-contigs Quality coverage: 6.85x in Q20 bases; agarose-fp

-----  
 \* NOTE: This is a 'working draft' sequence. It currently consists of 3 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

\* 1 194085: contig of 194085 bp in length  
 \* 194086 194185: gap of 100 bp  
 \* 194186 204837: contig of 10652 bp in length  
 \* 204838 204937: gap of 100 bp  
 \* 204938 327883: contig of 122946 bp in length.

## FEATURES

source

Location/Qualifiers  
 1..327883

/organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"

/chromosome="1"  
 /clone="CTD-3032124"

/clone\_lib="CIT-HSP-D2"  
 1..194085

/note="assembly\_fragment:02147"

fragment\_chain:1

194186..204837

/note="assembly\_fragment:01387"

fragment\_chain:1

204938..327883

/note="assembly\_fragment:02101"

## ORIGIN

Query Match 88.0%; Score 26.4; DB 14; Length 327883;  
 Best Local Similarity 96.4%; Pred. No. 6.9;  
 Matches 27; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGAATAAAGCGTGTGTTAAAGACAAA 28

Db 139484 CGGAATAAAGCGTGTGTTAAAGACAAA 139511

## RESULT 11

LOCUS CQ922970 509 bp DNA linear PAT 23-NOV-2004  
 DEFINITION Sequence 4170 from Patent WO2004097052.  
 ACCESSION CQ922970  
 VERSION CQ922970.1 GI:56212911  
 KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homidae; Homo.

## REFERENCE

1  
 Strachan, A., Immerman, P. and Dörner, A. J.  
 Methods for prognosis and treatment of solid tumors

JOURNAL Patent: WO 2004097052-A 4170 11-NOV-2004;  
 Wyeth (US); Burczynski, Michael E. (US)

## FEATURES

source

Location/Qualifiers  
 1..509

/organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"

100  
 /note="n includes a, c, g, or t, or contains no nucleotide"

110  
 /note="n includes a, c, g, or t, or contains no nucleotide"

163  
 /note="n includes a, c, g, or t, or contains no nucleotide"

misc\_feature

misc\_feature 185 /note="n includes a, c, g, or t, or contains no nucleotide"  
misc\_feature 257..260 /note="n includes a, c, g, or t, or contains no nucleotide"  
misc\_feature 300 /note="n includes a, c, g, or t, or contains no nucleotide"  
misc\_feature 307 /note="n includes a, c, g, or t, or contains no nucleotide"  
misc\_feature 348..367 /note="n includes a, c, g, or t, or contains no nucleotide"  
misc\_feature 422..483 /note="n includes a, c, g, or t, or contains no nucleotide"

ORIGIN

Query Match 81.3%; Score 24.4; DB 6; Length 509;  
Best Local Similarity 83.3%; Pred. No. 1.5e+02;  
Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGAAATAAGGCTGTTGTAAGACAAAAA 30  
|||||  
Db 236 CGGAAATAAGGCTGTTGTAANNNGAAAA 265

RESULT 12  
AP000876/c  
LOCUS AP000876 161151 bp DNA linear HTG 30-MAY-2000  
DEFINITION Homo sapiens chromosome 18 clone RP11-771B1 map 18p11.3, WORKING DRAFT SEQUENCE, 17 unordered pieces.  
ACCESSION AP000876  
VERSION AP000876.2 GI:8119024  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 161151)  
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.  
Homo sapiens 161,151 genomic DNA of 18p11.3  
Published Only in Database (1999)  
2 (bases 1 to 161151)  
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.  
Direct Submission  
Submitted (13-DEC-1999) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555, Japan (E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/, Tel:81-42-778-9923, Fax:81-42-778-9924)  
On May 31, 2000 this sequence version replaced gi:6997705.  
----- Genome Center  
Center: RIKEN Genomic Sciences Center(GSC)  
Center code: RIKEN  
Web site: http://hgp.gsc.riken.go.jp/  
Contact: hattori@gsc.riken.go.jp  
----- Project Information  
Center project name: Humdraft18  
Center clone name: RP11-771B1  
----- Summary Statistics  
Sequencing vector: PCR products; 100% of reads  
Chemistry: Dye-terminator ET-amersham; 100% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 147768 bases at least Q40

Consensus quality: 154700 bases at least Q30  
Consensus quality: 157838 bases at least Q20  
Insert size: 159551; sum-of-contigs  
Quality coverage: 4.74x in Q20 bases; sum-of-contigs  
-----  
NOTE: This is a 'working draft' sequence. It currently consists of 17 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved

1 24287 contig of 24287 bp in length  
24388 4446 contig of 20059 bp in length  
59040 contig of 14494 bp in length  
74881 contig of 15741 bp in length  
89942 contig of 14961 bp in length  
103712 contig of 13670 bp in length  
114497 contig of 10685 bp in length  
125844 contig of 11247 bp in length  
133701 contig of 7757 bp in length  
140062 contig of 6261 bp in length  
145470 contig of 5308 bp in length  
150374 contig of 4804 bp in length  
150475 154892 contig of 4418 bp in length  
154993 155990 contig of 998 bp in length  
156091 157891 contig of 1801 bp in length  
157892 159471 contig of 1480 bp in length  
159472 159571 contig of 1480 bp in length  
159572 161151 contig of 1580 bp in length  
Sequence updated (26-May-2000).

\* NOTE: This is a 'working draft' sequence. It currently consists of 17 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. \* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved

1 24287: contig of 24287 bp in length  
24387: gap of 100 bp  
24388 4446: contig of 20059 bp in length  
44447 44546: gap of 100 bp  
44547 59040: contig of 14494 bp in length  
59041 59140: gap of 100 bp  
59141 74881: contig of 15741 bp in length  
74882 74981: gap of 100 bp  
74982 89942: contig of 14961 bp in length  
89943 90042: gap of 100 bp  
90043 103712: contig of 13670 bp in length  
103713 103812: gap of 100 bp  
103813 114497: contig of 10685 bp in length  
114498 114597: gap of 100 bp  
114598 125844: contig of 11247 bp in length  
125845 125944: gap of 100 bp  
125945 133701: contig of 7757 bp in length  
133702 133801: gap of 100 bp  
133802 140062: contig of 6261 bp in length  
140063 140162: gap of 100 bp  
140163 145470: contig of 5308 bp in length  
145471 145571: gap of 100 bp  
145571 150374: contig of 4804 bp in length  
150375 150475: gap of 100 bp  
150475 154892: contig of 4418 bp in length  
154893 154993: contig of 998 bp in length  
154993 155990: gap of 100 bp  
155991 156091: contig of 1801 bp in length  
156091 157891: gap of 100 bp  
157892 159471: contig of 1480 bp in length  
159472 159571: gap of 100 bp  
159572 161151: contig of 1580 bp in length.  
Location/Qualifiers  
1. .161151

FEATURES  
source



```
repeat_region /rpt_family="MLT1I"
15540..15705
/rpt_family="MER5B"
15836..16582
/rpt_family="L1PA6"
complement(16583..17089)
/rpt_family="MER9"
17090..22509
/rpt_family="L1PA6"
22774..23406
/rpt_family="L1MEC"
complement(23983..24178)
/rpt_family="MLT1J"
complement(25899..26406)
/rpt_family="MLT1D"
26790..26810
/rpt_family="(CTGT)n"
complement(30033..31005)
/rpt_family="L2"
31546..31959
/rpt_family="MLT1D"
31960..32236
/rpt_family="MER44D"
complement(32237..32591)
/rpt_family="MER47A"
32592..32991
/rpt_family="MER44D"
complement(34134..34349)
/rpt_family="MIR"
34424..34847
/rpt_family="L1MC4"
complement(34871..34969)
/rpt_family="AluJo/FRAM"
35021..35144
/rpt_family="(TATATG)n"
35258..35358
/rpt_family="AT_rich"
35821..35658
/rpt_family="AT_rich"
complement(35783..35897)
/rpt_family="MIR"
complement(36089..36316)
/rpt_family="MIR"
36787..37087
/rpt_family="AluY"
37383..37556
/rpt_family="MER102"
37578..37607
/rpt_family="(CA)n"
39358..39383
/rpt_family="(T)n"
39936..40082
/rpt_family="MIR"
40495..40637
/rpt_family="MER5B"
41783..42073
/rpt_family="Alusg"
42856..42965
/rpt_family="MIR"
45120..45328
/rpt_family="L1MB8"
46246..46398
/rpt_family="MER5B"
complement(46520..46819)
/rpt_family="Alusq"
complement(47951..48228)
/rpt_family="MER33"
50241..50533
/rpt_family="AluJo"
51158..51215
/rpt_family="(TA)n"
51624..52514
/rpt_family="Tigger3b"
```

```
repeat_region 52515..52741
/rpt_family="MER30"
52742..52934
/rpt_family="Tigger3b"
53342..53375
/rpt_family="AT_rich"
53450..53754
/rpt_family="Alusx"
54957..55256
/rpt_family="AluYb8"
55888..55921
/rpt_family="AT_rich"
56168..56581
/rpt_family="MSTA"
56833..56920
/rpt_family="L1MA9"
57982..58106
/rpt_family="MER103"
58703..59203
/rpt_family="L1MC5"
complement(59212..59426)
/rpt_family="MER20"
59443..59650
/rpt_family="L1MC5"
59651..59937
/rpt_family="AluSp"
59938..60333
/rpt_family="L1MC5"
61866..61886
/rpt_family="AT_rich"
complement(62370..62523)
/rpt_family="AluJo"
62524..62557
/rpt_family="(TTTG)n"
```

Query Match 78.0%; Score 23.4; DB 8; Length 211945;  
Best Local Similarity 96.0%; Pred. No. 94;  
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 ATAAAGCTGTTGTTAAAGACAAAAA 30

DB 129603 ATAAAGCTGTTGTTAAAGATAAAAA 129627

## RESULT 14

BD275867/c  
LOCUS BD275867 301 bp DNA linear PAT 17-JUL-2003  
DEFINITION COMPOUNDS AND METHODS FOR THERAPY AND DIAGNOSIS OF LUNG CANCER.  
ACCESSION BD275867  
VERSION BD275867.1 GI:33085635  
KEYWORDS JP 2002543769-A/261.  
SOURCE unidentified  
ORGANISM unclassified.  
REFERENCE 1 (bases 1 to 301)  
AUTHORS Wang,T. and Van,R.  
TITLE COMPOUNDS AND METHODS FOR THERAPY AND DIAGNOSIS OF LUNG CANCER  
JOURNAL Patent: JP 2002543769-A 261 24-DEC-2002;  
Corixa Corporation et al  
COMMENT OS Homo sapien  
PN JP 2002543769-A/261  
PD 24-DEC-2002  
PF 03-APR-2000 JP 2000611554  
PR 22-FEB-2000 US 09/510376,10-JAN-2000 US 09/480884, PR  
02-APR-1999 US 09/285479,30-DEC-1999 US 09/476496, PR  
17-DEC-1999 US 09/466396  
PI tongtong wang,rikn van  
CC  
FH Key Location/Qualifiers  
FT misc\_feature (1)..(301)  
FT note='n = A,T,C or G'.  
FEATURES Location/Qualifiers  
source 1..301

/organism="unidentified"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32644"

ORIGIN

Query Match 76.7%; Score 23; DB 6; Length 301;  
Best Local Similarity 95.8%; Pred. No. 5.6e+02;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGAATAAAGGCTGTGTAAAGA 24  
|||  
Db 25 CGGAATAAAGGCTGTGTAAAGA 2

RESULT 15

AR220652/c  
LOCUS AR220652 301 bp DNA linear PAT 26-SEP-2002  
DEFINITION Sequence 304 from patent US 6426072.  
ACCESSION AR220652  
VERSION AR220652.1 GI:23327433  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
Unclassified.

REFERENCE 1 (bases 1 to 301)  
AUTHORS Wang, T., Fan, L., Kalos, M.D., Bangur, C.S., Hosken, N.A., Fanger, G.R.,  
Li, S.X., Wang, A., Skeiky, Y.A.W., Henderson, R.A. and McNeill, P.D.  
TITLE Compositions and methods for the therapy and diagnosis of lung  
cancer

JOURNAL Patent: US 6426072-A 304 30-JUL-2002;

Corixa Corporation; Seattle, WA  
Location/Qualifiers  
source 1. 301  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN

Query Match 76.7%; Score 23; DB 6; Length 301;  
Best Local Similarity 95.8%; Pred. No. 5.6e+02;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGAATAAAGGCTGTGTAAAGA 24  
|||  
Db 25 CGGAATAAAGGCTGTGTAAAGA 2

Search completed: December 7, 2005, 04:22:16  
Job time : 270.316 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 11:31:40 ; Search time 74.7009 Seconds  
(without alignments)  
3320.997 Million cell updates/sec

Title: US-09-980-046B-9  
Perfect score: 30  
Sequence: 1 cggaaataaaggctgttgaagacaaaaa 30

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA Main:  
1: /cgn2\_6/ptodata/1/pubna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubna/US08\_PUBCOMB.seq.\*  
3: /cgn2\_6/ptodata/1/pubna/US09A\_PUBCOMB.seq.\*  
4: /cgn2\_6/ptodata/1/pubna/US09B\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubna/US10A\_PUBCOMB.seq.\*  
6: /cgn2\_6/ptodata/1/pubna/US10B\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubna/US10C\_PUBCOMB.seq.\*  
8: /cgn2\_6/ptodata/1/pubna/US10D\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubna/US10E\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubna/US11\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	28.4	94.7	547	3	US-09-997-003-18
2	28.4	94.7	642	5	US-10-044-090-536
3	28.4	94.7	670	6	US-10-131-410-2
4	28.4	94.7	678	8	US-10-723-860-695
5	28.4	94.7	678	9	US-10-756-149-705
6	28.4	94.7	1323	6	US-10-264-049-307
7	28.4	94.7	1389	8	US-10-723-860-5322
8	23	76.7	301	3	US-09-735-703-304
9	23	76.7	301	3	US-09-850-716A-304
10	23	76.7	301	5	US-09-897-778-304
11	23	76.7	301	5	US-10-007-700-304
12	23	76.7	301	6	US-10-117-982-304
13	23	76.7	301	6	US-10-313-986-304
14	23	76.7	301	8	US-10-775-972-304
15	23	76.7	301	9	US-10-922-124-304
16	21.2	70.7	905	5	US-10-198-846-4040
17	21	70.0	464	3	US-09-764-887-1322
18	21	70.0	464	5	US-10-073-961-1322
19	21	70.0	486	3	US-09-764-887-586
20	21	70.0	486	3	US-09-764-887-587
21	21	70.0	486	5	US-10-073-961-586
22	21	70.0	486	5	US-10-073-961-587
23	21	70.0	634	3	US-09-954-531-144

24 21 70.0 634 3 US-09-954-531-362 Sequence 362, App  
25 21 70.0 634 3 US-09-954-531-571 Sequence 571, App  
26 21 70.0 634 6 US-10-177-293-233 Sequence 233, App  
27 21 70.0 634 6 US-10-101-510-190 Sequence 190, App  
28 21 70.0 634 6 US-10-159-563-140 Sequence 140, App  
29 21 70.0 634 7 US-10-240-425-1293 Sequence 1293, App  
30 21 70.0 634 7 US-10-188-832-105 Sequence 105, App  
31 21 70.0 634 7 US-10-755-889-231 Sequence 231, App  
32 21 70.0 634 8 US-10-775-169-214 Sequence 214, App  
33 21 70.0 634 8 US-10-802-432-4 Sequence 4, Appli  
34 21 70.0 634 8 US-10-370-715B-195 Sequence 195, App  
35 21 70.0 634 8 US-10-487-422-95 Sequence 95, Appl  
36 21 70.0 634 9 US-10-843-641A-1211 Sequence 1211, App  
37 21 70.0 634 9 US-10-843-641A-1429 Sequence 1429, App  
38 21 70.0 634 9 US-10-843-641A-1638 Sequence 1638, App  
39 21 70.0 634 9 US-10-631-467-67 Sequence 67, Appl  
C 40 20.6 68.7 148 8 US-10-674-124A-16152 Sequence 16152, A  
41 20.6 68.7 201 8 US-10-719-993-48862 Sequence 48862, A  
42 20.6 68.7 201 8 US-10-719-993-48904 Sequence 48904, A  
43 20.6 68.7 201 8 US-10-719-993-48920 Sequence 48920, A  
44 20.6 68.7 538 4 US-09-925-065A-739224 Sequence 739224,  
45 20.6 68.7 575 4 US-09-925-065A-732979 Sequence 732979,

## ALIGNMENTS

RESULT 1  
US-09-997-003-18  
; Sequence 18, Application US/09997003  
; Publication No. US20030203361A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PA003P1  
; CURRENT APPLICATION NUMBER: US/09/997,003  
; CURRENT FILING DATE: 2001-11-30  
; PRIOR APPLICATION NUMBER: unassigned  
; PRIOR FILING DATE: 2001-11-30  
; PRIOR APPLICATION NUMBER: PCT/US00/22157  
; PRIOR FILING DATE: 2000-08-11  
; PRIOR APPLICATION NUMBER: 60/148,680  
; PRIOR FILING DATE: 1999-08-13  
; NUMBER OF SEQ ID NOS: 56  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 18  
; LENGTH: 547  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-997-003-18

Query Match 94.7%; Score 28.4; DB 3; Length 547;  
Best Local Similarity 96.7%; Pred. No. 0.27; Indels 0; Gaps 0;  
Matches 29; Conservative 0; Mismatches 1;

QY 1 CGGAATAAAGCGCTGTTGTAAGACAAAAA 30  
|||||  
Db 488 CGGAATAAAGCGCTGTTGTAAGACAAAAA 517

RESULT 2  
US-10-044-090-536  
; Sequence 536, Application US/10044090  
; Publication No. US20020137081A1  
; GENERAL INFORMATION:  
; APPLICANT: Olga Bandman  
; TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION  
; FILE REFERENCE: PA-0028 US  
; CURRENT APPLICATION NUMBER: US/10/044,090  
; CURRENT FILING DATE: 2002-01-09  
; NUMBER OF SEQ ID NOS: 850  
; SOFTWARE: PERL Program  
; SEQ ID NO 536

```

; LENGTH: 642
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Inc
US-10-044-090-536

```

Query Match 94.7%; Score 28.4; DB 5; Length 642;  
Best Local Similarity 96.7%; Pred. No. 0.28;  
Matches 29; Conservative 0; Mismatches 1; Indels

Qy 1 CGGAATAAAGGCTGTTGTAAGACAAAA 30  
606 CGGAATAAAGGCTGTTGTAAGACAAAA 635

### RESULT 3

US-10-131-410-2  
Sequence 2, Application US/10131410  
; Publication No: US20030235915A1

; GENERAL INFORMATION:  
 ; APPLICANT: SPECHT, THOMAS  
 ; APPLICANT: HINZMANN, BERND  
 ; APPLICANT: SCHMITT, ARMIN  
 ; APPLICANT: PILARSKY, CHRISTIAN  
 ; APPLICANT: DAHL, EDGAR  
 ; APPLICANT: ROSENTHAL, ANDRE

; TITLE OF INVENTION: HUMAN NUCLEIC ACID SEQUENCES FROM TISSUE OF BREAST

TITLE OF INVENTION: TUMORS  
 FILE REFERENCE: SCH-1763  
 CURRENT APPLICATION NUMBER: US/10/131,410

; CURRENT FILING DATE: 2002-04-25  
; PRIOR APPLICATION NUMBER: 09/646,673

; PRIORITY FILING DATE: 2000-09-20  
 ; PRIORITY APPLICATION NUMBER: PCT/DE99/00

; PRIOR FILING DATE: 1999-03-19  
 ; NUMBER OF SEQ ID NOS: 202

```

; NUMBER OF SEQ ID NOS: 202
; SOFTWARE: PatentIn Ver. 2.1

```

```

; SEQ ID NO 2
; LENGTH: 670

```

```

;
; TYPE: DNA
;
; ORGANISM: Homo sapiens

```

US-10-131-410-2

Query Match 94.7

Best Local Similarity 96.7  
Matches 29; Conservative

OV 1 CGGAAATAAAGGCTGT

[illegible]

1

RESULT 4  
US-10-723-860-695

; Sequence 695, Application  
: Publication No. US200402

; APPLICANT: Aziz, Natasha  
 ; APPLICANT: Ginsburg, Wendy M.  
 ; APPLICANT: Zlotnik, Albert  
 ; TITLE OF INVENTION: Methods of diagnosis of

; FILE REFERENCE: 05882.0193.NPUS01

```

; CURRENT APPLICATION NUMBER: US/10/72
;
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2003-11-26

```

; PRIOR FILING DATE: 2002-11-26  
 ;  
 ; NUMBER OF SEQ ID NOS: 8393  
 ;

```

; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 695

```

LENGTH: 678



RESULT 7  
US-10-723-860-5322  
; Sequence 5322, Application US/10723860  
; Publication No. US20040253606A1  
; GENERAL INFORMATION:  
; APPLICANT: Azila, Natasha  
; APPLICANT: Ginsburg, Wendy M.  
; APPLICANT: Zlotnik, Albert  
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &  
; TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators  
; FILE REFERENCE: 05882.0193.NPUS01  
; CURRENT APPLICATION NUMBER: US/10/723,860  
; CURRENT FILING DATE: 2003-11-26  
; PRIOR APPLICATION NUMBER: 60/429,739  
; PRIOR FILING DATE: 2002-11-26  
; NUMBER OF SEQ ID NOS: 8393  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 5322  
; LENGTH: 1389  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-723-860-5322

Query Match 94.7%; Score 28.4; DB 8; Length 1389;  
Best Local Similarity 96.7%; Pred. No. 0.32;  
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGAATAAAGGCTGTTGTAAGACAAAA 30  
|||  
Db 554 CGGAATAAAGGCTGTTGTAAGAGAAAA 583

RESULT 8  
US-09-735-705-304/c  
; Sequence 304, Application US/09735705  
; Patent No. US20020052329A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Fan, Liqun  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Hosken, Nancy  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Li, Samuel X.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: McNeill, Patricia D.  
; APPLICANT: Fanger, Neil  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.455C14  
; CURRENT APPLICATION NUMBER: US/09/735,705  
; CURRENT FILING DATE: 2000-12-12  
; NUMBER OF SEQ ID NOS: 419  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 304  
; LENGTH: 301  
; TYPE: DNA  
; ORGANISM: Homo sapien  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)-(301)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-735-705-304

Query Match 76.7%; Score 23; DB 3; Length 301;  
Best Local Similarity 95.8%; Pred. No. 33;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGAATAAAGGCTGTTGTAAGA 24  
|||  
Db 25 CGGAATAAAGGCTGTTGTAAGA 2

RESULT 9  
US-09-850-716A-304/c  
; Sequence 304, Application US/09850716A  
; Patent No. US20020115139A1  
; GENERAL INFORMATION:  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: McNeill, Patricia D.  
; APPLICANT: Retter, Marc W.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.455C15  
; CURRENT APPLICATION NUMBER: US/09/850,716A  
; CURRENT FILING DATE: 2001-05-07  
; NUMBER OF SEQ ID NOS: 440  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 304  
; LENGTH: 301  
; TYPE: DNA  
; ORGANISM: Homo sapien  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)-(301)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-850-716A-304

Query Match 76.7%; Score 23; DB 3; Length 301;  
Best Local Similarity 95.8%; Pred. No. 33;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGAATAAAGGCTGTTGTAAGA 24  
|||  
Db 25 CGGAATAAAGGCTGTTGTAAGA 2

RESULT 10  
US-09-897-778-304/c  
; Sequence 304, Application US/09897778  
; Patent No. US20020147143A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Marnerakis, Margarita  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Carter, Barrick  
; APPLICANT: Watanabe, Yoshihiro  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Peckham, David W.  
; APPLICANT: Fanger, Neil  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.455C16  
; CURRENT APPLICATION NUMBER: US/09/897,778  
; CURRENT FILING DATE: 2001-06-28  
; NUMBER OF SEQ ID NOS: 467  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 304  
; LENGTH: 301  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 23, 104, 192  
; OTHER INFORMATION: n = A,T,C or G  
US-09-897-778-304

Query Match 76.7%; Score 23; DB 3; Length 301;  
Best Local Similarity 95.8%; Pred. No. 33;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGAATAAAGGCTGTTGTAAGA 24  
|||  
Db 25 CGGAATAAAGGCTGTTGTAAGA 24

```
Db      25  CGNAATAAAGGCTGTTGTTAAAGA 2

RESULT 11
US-10-007-700-304/c
; Sequence 304, Application US/10007700
; Publication No. US20030064947A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Li, Samuel X.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Fanger, Neil
; APPLICANT: Retter, Marc W.
; APPLICANT: Durham, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Peckman, David W.
; APPLICANT: Cai, Feng
; APPLICANT: Foy, Teresa M.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C17
; CURRENT APPLICATION NUMBER: US/10/007,700
; CURRENT FILING DATE: 2001-11-30
; NUMBER OF SEQ ID NOS: 469
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 304
; LENGTH: 301
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 23, 104, 192
; OTHER INFORMATION: n = A,T,C or G
US-10-007-700-304

Query Match      76.7%; Score 23; DB 5; Length 301;
Best Local Similarity 95.8%; Pred. No. 33;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1  CGGAATAAAGGCTGTTGTTAAAGA 24
Db      25  CGNAATAAAGGCTGTTGTTAAAGA 2

RESULT 12
US-10-117-982-304/c
; Sequence 304, Application US/10117982
; Publication No. US20030138438A1
; GENERAL INFORMATION:
; APPLICANT: Foy, Teresa M.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Mericle, Barbara
; APPLICANT: Spies, Gregory A.
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C18
; CURRENT APPLICATION NUMBER: US/10/117,982
; CURRENT FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 484
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 304
; LENGTH: 301
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 23, 104, 192
; OTHER INFORMATION: n = A,T,C or G
US-10-117-982-304

Query Match      76.7%; Score 23; DB 5; Length 301;
Best Local Similarity 95.8%; Pred. No. 33;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1  CGGAATAAAGGCTGTTGTTAAAGA 24
Db      25  CGNAATAAAGGCTGTTGTTAAAGA 2

RESULT 13
US-10-313-986-304/c
; Sequence 304, Application US/10313986
; Publication No. US20030236209A1
; GENERAL INFORMATION:
; APPLICANT: Foy, Teresa M.
; APPLICANT: McNabb, Andria
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Reed, Steven G.
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C19
; CURRENT APPLICATION NUMBER: US/10/313,986
; CURRENT FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 304
; LENGTH: 301
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 23, 104, 192
; OTHER INFORMATION: n = A,T,C or G
US-10-313-986-304

Query Match      76.7%; Score 23; DB 6; Length 301;
Best Local Similarity 95.8%; Pred. No. 33;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1  CGGAATAAAGGCTGTTGTTAAAGA 24
Db      25  CGNAATAAAGGCTGTTGTTAAAGA 2

RESULT 14
US-10-775-972-304/c
; Sequence 304, Application US/10775972
; Publication No. US20040235072A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C21
; CURRENT APPLICATION NUMBER: US/10/775,972
; CURRENT FILING DATE: 2004-02-10
; NUMBER OF SEQ ID NOS: 563
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 304
; LENGTH: 301
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 23, 104, 192
; OTHER INFORMATION: n = A,T,C or G
US-10-775-972-304/c

Query Match      76.7%; Score 23; DB 6; Length 301;
Best Local Similarity 95.8%; Pred. No. 33;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 23, 104, 192
; OTHER INFORMATION: n = A,T,C or G
US-10-775-972-304

Query Match          76.7%; Score 23; DB 8; Length 301;
Best Local Similarity 95.8%; Pred. No. 33;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGAATAAAGGCTGTCTTAAAGA 24
   |||||
Db 25 CGNAATAAAGGCTGTCTTAAAGA 2
```

```
RESULT 15
US-10-922-124-304/c
; Sequence 304, Application US/10922124
; Publication No. US20050142620A1
; GENERAL INFORMATION:
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Zehentner-Wilkinson, Barbara K.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C22
; CURRENT APPLICATION NUMBER: US/10/922,124
; CURRENT FILING DATE: 2004-08-18
; PRIOR APPLICATION NUMBER: 10/775,972
; PRIOR FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: 10/623,155
; PRIOR FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 10/313,986
; PRIOR FILING DATE: 2002-12-04
; PRIOR APPLICATION NUMBER: 10/117,982
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 10/007,700
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 09/897,778
; PRIOR FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: 09/850,716
; PRIOR FILING DATE: 2001-05-07
; PRIOR APPLICATION NUMBER: 09/735,705
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: 09/685,696
; PRIOR FILING DATE: 2000-10-09
; PRIOR APPLICATION NUMBER: 09/662,786
; PRIOR FILING DATE: 2000-09-15
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 563
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 304
; LENGTH: 301
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 23, 104, 192
; OTHER INFORMATION: n = A,T,C or G
US-10-922-124-304
```

```
Query Match          76.7%; Score 23; DB 9; Length 301;
Best Local Similarity 95.8%; Pred. No. 33;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGAATAAAGGCTGTCTTAAAGA 24
   |||||
Db 25 CGNAATAAAGGCTGTCTTAAAGA 2
```

Search completed: December 6, 2005, 22:57:12  
Job time : 76.7009 secs

*This Page Blank (Lap 1)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 21:43:57 ; Search time 970.211 Seconds  
(without alignments)  
4099.008 Million cell updates/sec

Title: US-09-980-046B-10  
Perfect score: 85  
Sequence: 1 cggfagaagtcattcttcgc.....tgccccactaagttagaaaaa 85

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues  
Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
EST:  
1: gb\_est1:  
2: gb\_est2:  
3: gb\_est3:  
4: gb\_hc:  
5: gb\_est4:  
6: gb\_est5:  
7: gb\_est6:  
8: gb\_est7:  
9: gb\_g861:  
10: gb\_g862:  
11: gb\_g863:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	83.4	98.1	155	AA235328	AA235328
C 2	83.4	98.1	186	AA292318	AA292318
C 3	83.4	98.1	243	AA631632	AA631632
C 4	83.4	98.1	248	BF590563	BF590563
C 5	83.4	98.1	268	AA514303	AA514303
C 6	83.4	98.1	270	AI334968	AI334968
C 7	83.4	98.1	272	AI307267	AI307267
C 8	83.4	98.1	272	BF437551	BF437551
C 9	83.4	98.1	273	AI400926	AI400926
C 10	83.4	98.1	273	AI669789	AI669789
C 11	83.4	98.1	300	AI718891	AI718891
C 12	83.4	98.1	303	BE646272	BE646272
C 13	83.4	98.1	304	AI417902	AI417902
C 14	83.4	98.1	309	AI760920	AI760920
C 15	83.4	98.1	309	AA148966	AA148966
C 16	83.4	98.1	315	AI470973	AI470973
C 17	83.4	98.1	321	AI468348	AI468348
C 18	83.4	98.1	325	CK430031	CK430031
C 19	83.4	98.1	327	AI867674	AI867674
C 20	83.4	98.1	329	BF063247	BF063247
C 21	83.4	98.1	336	AI708880	AI708880
C 22	83.4	98.1	338	AI379991	AI379991

C 23	83.4	98.1	354	1	AW768412
C 24	83.4	98.1	355	1	AW627370
C 25	83.4	98.1	362	1	AA974311
C 26	83.4	98.1	365	2	BF059613
C 27	83.4	98.1	365	8	RI2367
C 28	83.4	98.1	379	5	BU963341
C 29	83.4	98.1	384	1	AA938338
C 30	83.4	98.1	387	6	CA436188
C 31	83.4	98.1	391	1	AI673220
C 32	83.4	98.1	395	1	AI659337
C 33	83.4	98.1	397	1	AI368956
C 34	83.4	98.1	402	3	BQ011944
C 35	83.4	98.1	405	1	AW452170
C 36	83.4	98.1	417	1	AI819871
C 37	83.4	98.1	420	1	AI028759
C 38	83.4	98.1	423	1	AI673198
C 39	83.4	98.1	429	6	CF529461
C 40	83.4	98.1	431	2	BF003132
C 41	83.4	98.1	443	1	AI870930
C 42	83.4	98.1	443	3	BQ011066
C 43	83.4	98.1	445	1	AI278959
C 44	83.4	98.1	448	1	AA100479
C 45	83.4	98.1	448	8	W74705

ALIGNMENTS

RESULT 1  
AA235328/c

LOCUS

DEFINITION

IMAGE:723854 3' similar to gb:L11045\_rna1 PROTEASOME COMPONENT C13

(HUMAN); mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

COMMENT

FEATURES

source

1.155

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:723854"

/sex="Female"

/tissue\_type="ovarian tumor"

/lab\_host="DH10B (ampicillin resistant)"

/clone\_lib="Soares ovary tumor NbHOT"

AA235328 155 bp mRNA linear EST 08-AUG-1997  
zt30c08.s1 Soares ovary tumor NbHOT Homo sapiens cDNA clone  
IMAGE:723854 3' similar to gb:L11045\_rna1 PROTEASOME COMPONENT C13  
(HUMAN); mRNA sequence.

AA235328 1 GI:1859965  
EST.  
Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
1 (Bases 1 to 155)  
Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B.,  
Chisoe, S., Dietrich, N., Dubucq, T., Favello, A., Gish, W.,  
Hawkins, M., Kuhlman, M., Kucaba, T., Lacy, M., Le, M., Le, N.,  
Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L.,  
Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J.,  
Trevas, E., Underwood, K., Wohldmann, P., Waterston, R., Wilson, R.  
and Marra, M.

Generation and analysis of 280,000 human expressed sequence tags  
Genome Res. 6 (9), 807-828 (1996)  
8889549  
Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Insert Length: 254 Std Error: 0.00  
Seq primer: -41m13 fwd. ET from Amerham.

Location/Qualifiers

1.155

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:723854"

/sex="Female"

/tissue\_type="ovarian tumor"

/lab\_host="DH10B (ampicillin resistant)"

/clone\_lib="Soares ovary tumor NbHOT"

/note="Organ: ovary; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site\_1: Not 1; Site\_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCAATCTGAAGTGGAGCGCGGTTTTTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library constructed by Bento Soares and M.Fatima Bonaldo."

ORIGIN

Query Match 98.1%; Score 83.4; DB 1; Length 155;  
Best Local Similarity 98.8%; Pred. No. 3.5e-18;  
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGTGAAGTGCATCTTCGCGTGTCTCTATTGAACAAGCATTTCCCGGAGGAACTTT 60  
Db 99 CGGTGAAGTGCATCTTCGCGTGTCTCTATTGAACAAGCATTTCCCGGAGGAACTTT 40

Qy 61 CTGGGTGCCCCCACTAAGTAGAAAA 85  
Db 39 CTGGGTGCCCCCACTAAGTAGAATAA 15

RESULT 2  
AA292318  
LOCUS  
DEFINITION  
z30c08.r1 Soares ovary tumor N8HOT Homo sapiens cDNA clone  
IMAGE:723854 5' similar to gb:U11045\_rnal PROTEASOME COMPONENT C13  
(HUMAN); mRNA sequence.

ACCESSION  
AA292318.1 GI:1940501  
VERSION  
AA292318  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.

REFERENCE  
1 (bases 1 to 186)  
Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B.,  
Chlesoe, S., Dietrich, N., Dubuque, T., Favell, A., Gish, W.,  
Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N.,  
Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L.,  
Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J.,  
Trevisan, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R.  
and Marra, M.

TITLE  
JOURNAL  
PUBMED  
COMMENT  
Generation and analysis of 280,000 human expressed sequence tags  
Genome Res. 6 (9), 807-828 (1996)  
Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810

Email: est@watson.wustl.edu  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.lnl.gov) for further information.  
Putative full length read  
The vector to vector length is 187  
Insert Length: 254 Std Error: 0.00  
Seq primer: -28ml3 rev2 ET from Amersham.

FEATURES  
Location/Qualifiers  
1..186  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:723854"  
/sex="Female"  
/tissue\_type="ovarian tumor"  
/lab\_host="DH10B (ampicillin resistant)"  
/clone\_lib="Soares ovary tumor N8HOT"  
/note="Organ: ovary; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site\_1: Not 1; Site\_2: Eco RI; 1st

strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCAATCTGAAGTGGAGCGCGGTTTTTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library constructed by Bento Soares and M.Fatima Bonaldo."

ORIGIN

Query Match 98.1%; Score 83.4; DB 1; Length 186;  
Best Local Similarity 98.8%; Pred. No. 3.6e-18;  
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGTGAAGTGCATCTTCGCGTGTCTCTATTGAACAAGCATTTCCCGGAGGAACTTT 60  
Db 89 CGGTGAAGTGCATCTTCGCGTGTCTCTATTGAACAAGCATTTCCCGGAGGAACTTT 148

Qy 61 CTGGGTGCCCCCACTAAGTAGAAAA 85  
Db 149 CTGGGTGCCCCCACTAAGTAGAATAA 173

RESULT 3  
AA631632/c  
LOCUS  
DEFINITION  
np59a05.s1 NCI CGAP Br2 Homo sapiens cDNA clone IMAGE:1130624 3'  
similar to gb:U11045\_rnal PROTEASOME COMPONENT C13 (HUMAN); mRNA  
sequence.

ACCESSION  
AA631632  
VERSION  
AA631632.1 GI:2554243  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.

REFERENCE  
1 (bases 1 to 243)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: csaphe@mail.nih.gov

Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.  
Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: M. Bento Soares, Ph.D.  
cDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
[www-bio.lnl.gov/bbrp/image/image.html](http://www-bio.lnl.gov/bbrp/image/image.html)  
Insert Length: 651 Std Error: 0.00  
Seq primer: -40ml3 fwd. ET from Amersham.

FEATURES  
Location/Qualifiers  
1..243  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:1130624"  
/sex="female, pooled"  
/tissue\_type="breast"  
/lab\_host="DH10B"  
/clone\_lib="NCI-CGAP Br2"  
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from pooled bulk  
breast tumor tissue, and was then primed with a Not I -  
oligo(dT) primer. Double-stranded cDNA was ligated to Eco  
RI adaptors (Pharmacia), digested with Not I and cloned  
into the Not I and Eco RI sites of the modified pT7T3  
vector. This library is the normalized version of  
NCI CGAP Br1.1. Library was constructed by Bento Soares  
and M. Fatima Bonaldo."

ORIGIN

Query Match 98.1%; Score 83.4; DB 1; Length 243;  
 Best Local Similarity 98.8%; Pred. No. 3.9e-18;  
 Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGTGAAGTGCATCTCTCGGTGTTCTCTATTGAAACAGCATTTCCCGCCAGGGAAGTTT 60  
 DB 103 CGGTGAAGTGCATCTCTCGGTGTTCTCTATTGAAACAGCATTTCCCGCCAGGGAAGTTT 44

QY 61 CTGGGTGCCCACTAAGTAGAATAA 85

DB 43 CTGGGTGCCCACTAAGTAGAATAA 19

RESULT 4  
 BF590563/c  
 LOCUS  
 DEFINITION 7h38e05.x1 NCI\_CGAP\_Col6 Homo sapiens cDNA clone IMAGE:3318272 3', mRNA sequence.

ACCESSION BF590563  
 VERSION BF590563.1 GI:11682887

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

REFERENCE 1 (bases 1 to 248)  
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: Ilan Kirsch, M.D., Michael R. Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D.  
 cDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL, send email to: info@image.llnl.gov

Seq primer: -40UP from Gibco.

Location/Qualifiers

1. .248

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:3318272"

/tissue\_type="colon tumor, RER+"

/lab\_host="DH10B"

/clone\_lib="NCI CGAP Col6"

/note="Organ: colon; Vector: p7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not 1; Site 2: Eco RI; Plasmid DNA from the normalized library NCI\_CGAP Col0 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIds 1057416-1061255, and 1144584-1145351). Subtraction by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 98.1%; Score 83.4; DB 2; Length 248;

Best Local Similarity 98.8%; Pred. No. 3.9e-18;

Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGTGAAGTGCATCTCTCGGTGTTCTCTATTGAAACAGCATTTCCCGCCAGGGAAGTTT 60

DB 104 CGGTGAAGTGCATCTCTCGGTGTTCTCTATTGAAACAGCATTTCCCGCCAGGGAAGTTT 45

QY 61 CTGGGTGCCCACTAAGTAGAATAA 85

|||||

DB 44 CTGGGTGCCCACTAAGTAGAATAA 20

RESULT 5

RA514303/c

LOCUS

DEFINITION

AA514303

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Elias Campo, M.D., Michael R. Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D.

cDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 354 Std Error: 0.00

Seq primer: -40ml3 fwd. ET from Amersham.

Location/Qualifiers

1. .268

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:924306"

/sex="pooled"

/tissue\_type="colon"

/lab\_host="DH10B"

/clone\_lib="NCI CGAP Co3"

/note="Vector: p7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not 1; Site 2: Eco RI; 1st strand cDNA was prepared from 12 pooled bulk tumor samples and primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p7T3 vector. Library went through one round of normalization."

ORIGIN

Query Match 98.1%; Score 83.4; DB 1; Length 268;

Best Local Similarity 98.8%; Pred. No. 3.9e-18;

Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGTGAAGTGCATCTCTCGGTGTTCTCTATTGAAACAGCATTTCCCGCCAGGGAAGTTT 60

DB 100 CGGTGAAGTGCATCTCTCGGTGTTCTCTATTGAAACAGCATTTCCCGCCAGGGAAGTTT 41

QY 61 CTGGGTGCCCACTAAGTAGAATAA 85

DB 40 CTGGGTGCCCACTAAGTAGAATAA 16

|||||

RESULT 6

AI334968/c

LOCUS

DEFINITION

AI334968

tb2ib05.x1 NCI CGAP\_Kid12 Homo sapiens cDNA clone IMAGE:2054961 3', similar to gb:L11045\_rnal PROTEASOME COMPONENT C13 (HUMAN); mRNA sequence.

270 bp mRNA linear EST 16-FEB-1999

```

ACCESSION      AI334968
VERSION         AI334968.1  GI:4071895
KEYWORDS        EST.
SOURCE          Homo sapiens (human)
ORGANISM        Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE      1 (bases 1 to 270)
AUTHORS        NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE          National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
JOURNAL        Tumor Gene Index
COMMENT        Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapsb@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
DNA Sequencing by: Greg Lennon, Ph.D.
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 355 Std Error: 0.00
Seq primer: -40UP from Gibco.
Location/Qualifiers
FEATURES       1..270
source         /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /clone="IMAGE:2054961"
               /tissue_type="2 pooled tumors (clear cell type)"
               /lab_host="DH10B"
               /clone_lib="NCI CGAP Kid12"
               /note="Organ: Kidney; Vector: pT7T3D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not I; Site 2: Eco RI;
Plasmid DNA from the normalized library NCI CGAP Kid5 was
prepared, and ss circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(cloneIDs 1323912-1325831, 1471368-1472903 and
1492104-1493255). Subtraction by Bento Soares and M.
Fatima Bonaldo. "
ORIGIN
Query Match      98.1%; Score 83.4; DB 1; Length 270;
Best Local Similarity 98.8%; Pred. No. 3.9e-18;
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 CGGTGAAGTGCATCTCTCGGTGTTCTCTATTGAACAAGCATTTCCCCAGGGAAGTTT 60
Db 110 CGGTGAAGTGCATCTCTCGGTGTTCTCTATTGAACAAGCATTTCCCCAGGGAAGTTT 51
Qy 61 CTGGTGCCCACTAAGTAGAAAAA 85
Db 50 CTGGTGCCCACTAAGTAGAATAA 26
RESULT 7
AI307267/c
LOCUS          tb24d10.x1 NCI CGAP Kid12 Homo sapiens cDNA clone IMAGE:205283 3'
DEFINITION     similar to gb:L11045_rna1 PROTEASOME COMPONENT C13 (HUMAN); mRNA
sequence.
ACCESSION      AI307267
VERSION        AI307267.1  GI:4002456
KEYWORDS        EST.
SOURCE          Homo sapiens (human)
ORGANISM        Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE      1 (bases 1 to 270)
AUTHORS        NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE          National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
JOURNAL        Tumor Gene Index
COMMENT        Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapsb@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
DNA Sequencing by: Greg Lennon, Ph.D.
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 355 Std Error: 0.00
Seq primer: -40UP from Gibco.
Location/Qualifiers
FEATURES       1..270
source         /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /clone="IMAGE:2054961"
               /tissue_type="2 pooled tumors (clear cell type)"
               /lab_host="DH10B"
               /clone_lib="NCI CGAP Kid12"
               /note="Organ: Kidney; Vector: pT7T3D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not I; Site 2: Eco RI;
Plasmid DNA from the normalized library NCI CGAP Kid5 was
prepared, and ss circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(cloneIDs 1323912-1325831, 1471368-1472903 and
1492104-1493255). Subtraction by Bento Soares and M.
Fatima Bonaldo. "
ORIGIN
Query Match      98.1%; Score 83.4; DB 1; Length 270;
Best Local Similarity 98.8%; Pred. No. 3.9e-18;
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 CGGTGAAGTGCATCTCTCGGTGTTCTCTATTGAACAAGCATTTCCCCAGGGAAGTTT 60
Db 110 CGGTGAAGTGCATCTCTCGGTGTTCTCTATTGAACAAGCATTTCCCCAGGGAAGTTT 51
Qy 61 CTGGTGCCCACTAAGTAGAAAAA 85
Db 50 CTGGTGCCCACTAAGTAGAATAA 26
RESULT 8
BF437551/c
LOCUS          7p74a08.x1 NCI CGAP P-28 Homo sapiens cDNA clone IMAGE:3651374 3',
DEFINITION     mRNA sequence.
ACCESSION      BF437551
VERSION        BF437551.1  GI:11449957
KEYWORDS        EST.
SOURCE          Homo sapiens (human)
ORGANISM        Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE      1 (bases 1 to 272)
AUTHORS        NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE          National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
JOURNAL        Tumor Gene Index
COMMENT        Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapsb@mail.nih.gov
Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.

```



cDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL, send email to:  
 info@image.llnl.gov

Seq primer: -40UP from Gibco.

#### FEATURES

Location/Qualifiers  
 1. .272  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3651374"  
 /sex="male"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 /clone\_lib="NCI CGAP Pr28"  
 /note="Organ: prostate; Vector: p773D-Pac (Pharmacia)  
 with a modified polylinker; Plasmid DNA from the  
 normalized library NCI CGAP Pr22 was prepared, and ss  
 circles were made in vitro. Following HAP purification,  
 this DNA was used as tracer in a subtractive hybridization  
 reaction. The driver was PCR-amplified cDNAs from a pool  
 of 5,000 clones made from the same library (cloneIDs  
 985608-986759, 110192-1101959, and 1217928-1220615).  
 Subtraction by Bento Soares and M. Fatima Bonaldo."

#### ORIGIN

Query Match 98.1%; Score 83.4; DB 2; Length 272;  
 Best Local Similarity 98.8%; Pred. No. 4e-18; Indels 0; Gaps 0;  
 Matches 84; Conservative 0; Mismatches 1;

QY 1 CGGTGAAGTGCATCTTCTCGGTGTTCTATTGAAACAGCATTTCCCCAGGGAAGTTT 60  
 |||||  
 Db 104 CGGTGAAGTGCATCTTCTCGGTGTTCTATTGAAACAGCATTTCCCCAGGGAAGTTT 45

QY 61 CTGGGTGCCCACTAAGTAGAATAA 85  
 |||||

Db 44 CTGGGTGCCCACTAAGTAGAATAA 20  
 |||||

#### RESULT 9

AI400926/c  
 LOCUS 273 bp mRNA linear EST 30-MAR-1999  
 th25e10.x1 NCI CGAP Pr28 Homo sapiens cDNA clone IMAGE:2119338 3'  
 DEFINITION similar to gb:L11045\_rnal PROTEASOME COMPONENT C13 (HUMAN); mRNA  
 sequence.

ACCESSION AI400926 GI:4244013

VERSION EST.

KEYWORDS Homo sapiens (human)

#### SOURCE

ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Hominidae; Homo.

REFERENCE 1 (bases 1 to 273)

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

Tumor Gene Index

Unpublished (1997)

CONTACT: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.

Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D.

DNA Sequencing by: Greg Lennon, Ph.D.

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 348 Std Error: 0.00

Seq primer: -40UP from Gibco.

Location/Qualifiers

1. .273

#### FEATURES

Location/Qualifiers

1. .273

/organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:2119338"  
 /sex="male"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"

/clone\_lib="NCI CGAP Pr28"  
 /note="Organ: prostate; Vector: p773D-Pac (Pharmacia)  
 with a modified polylinker; Plasmid DNA from the  
 normalized library NCI CGAP Pr22 was prepared, and ss  
 circles were made in vitro. Following HAP purification,  
 this DNA was used as tracer in a subtractive hybridization  
 reaction. The driver was PCR-amplified cDNAs from a pool  
 of 5,000 clones made from the same library (cloneIDs  
 985608-986759, 110192-1101959, and 1217928-1220615).  
 Subtraction by Bento Soares and M. Fatima Bonaldo."

#### ORIGIN

Query Match 98.1%; Score 83.4; DB 1; Length 273;  
 Best Local Similarity 98.8%; Pred. No. 4e-18; Indels 0;  
 Matches 84; Conservative 0; Mismatches 1; Gaps 0;

QY 1 CGGTGAAGTGCATCTTCTCGGTGTTCTATTGAAACAGCATTTCCCCAGGGAAGTTT 60  
 |||||  
 Db 105 CGGTGAAGTGCATCTTCTCGGTGTTCTATTGAAACAGCATTTCCCCAGGGAAGTTT 46

QY 61 CTGGGTGCCCACTAAGTAGAATAA 85  
 |||||

Db 45 CTGGGTGCCCACTAAGTAGAATAA 21  
 |||||

#### RESULT 10

AI669789/c

LOCUS 273 bp mRNA linear EST 15-DEC-1999

tusjic11.x1 NCI CGAP Pr28 Homo sapiens cDNA clone IMAGE:2252660 3'

DEFINITION similar to gb:L11045\_rnal PROTEASOME COMPONENT C13 (HUMAN); mRNA

sequence.

ACCESSION AI669789

VERSION AI669789.1 GI:4834563

KEYWORDS EST.

SOURCE Homo sapiens (human)

#### ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Hominidae; Homo.

REFERENCE 1 (bases 1 to 273)

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

CONTACT: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.

Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D.

DNA Sequencing by: Greg Lennon, Ph.D.

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 379 Std Error: 0.00

Seq primer: -40UP from Gibco.

Location/Qualifiers

1. .273

#### FEATURES

Location/Qualifiers

1. .273

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:2252660"

/sex="male"

/dev\_stage="adult"

/lab\_host="DH10B"

/clone\_lib="NCI CGAP Pr28"



```

QY 61 CTGGTGCCCACTAAGTAGAAAA 85
|||||
Db 39 CTGGTGCCCACTAAGTAGAATA 15

RESULT 13
AI417902/c
LOCUS
DEFINITION
t955b11.x1 NCI CGAP Pr28 mRNA linear EST 30-MAR-1999
similar to gb:L11045_rnal PROTEASOME COMPONENT C13 (HUMAN); mRNA
sequence.
ACCESSION
VERSION AI417902.1 GI:4261406
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 304)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 334 Std Error: 0.00
Seq primer: -40UP from Gibco.
Location/Qualifiers
1..304
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2112669"
/sex="male"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="NCI CGAP Pr28"
/note="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia)
with a modified polylinker; plasmid DNA from the
normalized library NCI CGAP Pr22 was prepared, and ss
circles were made in vitro. Following HAP purification,
this DNA was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from a pool
of 5,000 clones made from the same library (cloneids
985608-986759, 1101192-1101959, and 1217928-1220615).
Subtraction by Bento Soares and M. Fatima Bonaldo."

ORIGIN
Query Match 98.1%; Score 83.4; DB 1; Length 304;
Best Local Similarity 98.8%; Pred. No. 4.1e-18;
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGTGAAGTGCATCTCTCGGTGTTCTATTGAAACAGCATTTCCCCAGGGAAGTTT 60
|||||
Db 100 CGGTGAAGTGCATCTCTCGGTGTTCTATTGAAACAGCATTTCCCCAGGGAAGTTT 41

QY 61 CTGGTGCCCACTAAGTAGAAAA 85
|||||
Db 40 CTGGTGCCCACTAAGTAGAATA 16

RESULT 14
AI760920/c
LOCUS
DEFINITION
w170b11.x1 NCI CGAP Kid12 Homo sapiens cDNA clone IMAGE:2398653 3',
similar to gb:L11045_rnal PROTEASOME COMPONENT C13 (HUMAN); mRNA
sequence.
ACCESSION
VERSION AI760920.1 GI:5176587
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 306)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 387 Std Error: 0.00
Seq primer: -40UP from Gibco.
Location/Qualifiers
1..306
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2398653"
/tissue_type="2 pooled tumors (clear cell type)"
/lab_host="DH10B"
/clone_lib="NCI CGAP Kid12"
/note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not 1; Site 2: Eco RI;
Plasmid DNA from the normalized library NCI CGAP Kid5 was
prepared, and ss circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(cloneids 1323912-1325831, 1471368-1472903 and
1492104-1493255). Subtraction by Bento Soares and M.
Fatima Bonaldo."

ORIGIN
Query Match 98.1%; Score 83.4; DB 1; Length 306;
Best Local Similarity 98.8%; Pred. No. 4.1e-18;
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGTGAAGTGCATCTCTCGGTGTTCTATTGAAACAGCATTTCCCCAGGGAAGTTT 60
|||||
Db 112 CGGTGAAGTGCATCTCTCGGTGTTCTATTGAAACAGCATTTCCCCAGGGAAGTTT 53

QY 61 CTGGTGCCCACTAAGTAGAAAA 85
|||||
Db 52 CTGGTGCCCACTAAGTAGAATA 28

RESULT 15
AA148966/c
LOCUS
DEFINITION
zn99g11.s1 Stratagene colon (#937204) Homo sapiens cDNA clone
IMAGE:566372 3', similar to gb:L11045_rnal PROTEASOME COMPONENT C13
(HUMAN); mRNA sequence.
ACCESSION
VERSION AA148966.1 GI:1719064
KEYWORDS EST.
SOURCE Homo sapiens (human)

```

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.

REFERENCE  
AUTHORS 1 (bases 1 to 309)  
Hallier, L., Lennon, G., Becker, M., Bonaudo, M.F., Chiapelli, B.,  
Chisoe, S., Dietrich, N., DuBuque, T., Favello, A., Gish, W.,  
Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N.,  
Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L.,  
Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J.,  
Trevaskie, E., Underwood, K., Wohldmann, P., Waterston, R., Wilson, R.  
and Marra, M.

TITLE Generation and analysis of 280,000 human expressed sequence tags  
JOURNAL Genome Res. 6 (9), 807-828 (1996)  
PUBMED 8899549  
COMMENT Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu

FEATURES  
source 1..309  
    /organism="Homo sapiens"  
    /mol\_type="mRNA"  
    /db\_xref="GDB:4590825"  
    /db\_xref="taxon:9606"  
    /clone="IMAGE:566372"  
    /tissue="tumor"  
    /cell\_line="T84 carcinoma cell line"  
    /lab\_host="SOLR cells (kanamycin resistant)"  
    /clone\_lib="Stratagene colon (#937204)"  
    /notes="Organ: colon; Vector: pBluescript SK-; Site 1:  
EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:  
Oligo dT. T-84 colonic epithelial cell line. Average  
insert size: 1.0 kb; Uni-ZAP XR Vector; ~5' adaptor  
sequence: 5' GAATTCGACGAG 3' ~3' adaptor sequence: 5'  
CTCGAGTTTTTTTTTTTTTTT 3'"

ORIGIN

Query Match 98.1%; Score 83.4; DB 1; Length 309;  
Best Local Similarity 98.8%; Pred. No. 4.1e-18;  
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGTGAAGTGCATCTTCGCGTGTCTCTATTGCAACAGCATTTCCCCAGGGAGTTT 60  
    |||||  
Db 103 CGGTGAAGTGCATCTTCGCGTGTCTCTATTGCAACAGCATTTCCCCAGGGAGTTT 44  
    |||||

Qy 61 CTGGGTGCCCCACTAAGTAGAAAAA 85  
    |||||  
Db 43 CTGGGTGCCCCACTAAGTAGAATAA 19  
    |||||

Search completed: December 7, 2005, 09:56:06  
Job time : 973.311 secs

Result No.	Score	Query Match	Length	DB	ID	Description
1	93.4	25.2	79835	3	US-09-949-016-12456	Sequence 12456, A
2	93.4	25.2	79835	3	US-09-949-016-16121	Sequence 16121, A
c	90	24.3	9527	3	US-09-949-016-13979	Sequence 13979, A
4	89.4	24.2	20481	3	US-09-949-016-12093	Sequence 12093, A
5	89.4	24.2	20482	3	US-09-949-016-13660	Sequence 13660, A
c	88.6	23.9	601	3	US-09-949-016-143891	Sequence 143891, A
7	88.6	23.9	601	3	US-09-949-016-143892	Sequence 143892, A
8	88.6	23.9	601	3	US-09-949-016-143893	Sequence 143893, A
c	88.6	23.9	601	3	US-09-949-016-143894	Sequence 143894, A
10	88.6	23.9	601	3	US-09-949-016-143895	Sequence 143895, A
c	88.6	23.9	15330	3	US-09-949-016-15808	Sequence 15808, A
12	87.6	23.7	222452	3	US-09-949-016-12368	Sequence 12368, A
c	87	23.5	152070	3	US-09-949-016-15402	Sequence 15402, A
13	86.6	23.4	31444	3	US-09-949-016-17532	Sequence 17532, A
14	86.2	23.3	25260	3	US-09-949-016-11985	Sequence 11985, A
15	86.2	23.3	25260	3	US-09-949-016-12907	Sequence 12907, A
c	86	23.2	63783	3	US-09-949-016-12907	Sequence 12907, A
17	86	23.2	63783	3	US-09-949-016-13576	Sequence 13576, A
18	85.8	23.2	601	3	US-09-949-016-64807	Sequence 64807, A
19	85.8	23.2	21590	3	US-09-949-016-13372	Sequence 13372, A
20	85.8	23.2	21590	3	US-09-949-016-13373	Sequence 13373, A
21	85.6	23.1	28819	3	US-09-949-016-15806	Sequence 15806, A
22	85.4	23.1	61124	3	US-09-949-016-11914	Sequence 11914, A
23	85.4	23.1	61140	3	US-09-949-016-15771	Sequence 15771, A
24	84.8	22.9	601	3	US-09-949-016-42374	Sequence 42374, A



```
Query Match      24.2%; Score 89.4; DB 3; Length 20482;
Best Local Similarity 75.5%; Pred. No. 1.5e-19;
Matches 111; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 70 AGGTCATGCATCTGAGGTTTCATCAAAACATGGAAGCCATGCTAGTTATTTGTTCTTCAT 129
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 12360 AGGCCACATGCTGAGGTTTCATGAAGCACTGGAACCATGGTAGTTACTGTTCTTGAT 12419

QY 130 AGGTATGTCAACCTCAACACCAATGCTCTCTGCTGATGAGAAATGGGCTGCCCGAT 189
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 12420 AAGTATGTGGGCTCAACACCAAGTGGCTCTCTGCTGATGAAGAAATGGATGGCACTT 12479

QY 190 TCACAGACTCTTTGATCTTTCTCTCAGC 216
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 12480 GCCCTGAAAGGATGTCAGCTGCCCAGC 12506

RESULT 6
US-09-949-016-143891/c
; Sequence 143891, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 143891
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-143891

Query Match      23.9%; Score 88.6; DB 3; Length 601;
Best Local Similarity 72.3%; Pred. No. 4.9e-20;
Matches 115; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 212 TCAGCATATGATCTCTCATATGAACCTCTCTCTGCTAGGAGTGCGAGCCTACATCG 271
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 556 TCTCCTTTTCTCCACTTCTGACAGATCTTAGTCTTTGGCCAGGCGTGGTGCTCACATCT 497

QY 272 CTAATCCCGACCTTTGGGAAGCTGAGCGAGGAGTACTCTGAGGCCAGGAGTTTGAGA 331
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 496 ATAATCTAGCACTTTGGGAGGCTGAGGAGGAGGATCACTTGAATCCAGGAGTTTGAGA 437

QY 332 CCAGGCTGGGCAATACAGGAGACTCTCTCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCT 370
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 436 CCAGTCTGGGCAACACAGAGATCTCTCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 398

RESULT 7
US-09-949-016-143892/c
; Sequence 143892, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
```

```
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 143892
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-143892

Query Match      23.9%; Score 88.6; DB 3; Length 601;
Best Local Similarity 72.3%; Pred. No. 4.9e-20;
Matches 115; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 212 TCAGCATATGATCTCTCATATGAACCTCTCTCTGCTAGGAGTGCGAGCCTACATCG 271
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 544 TCTCCTTTTCTCCACTTCTGACAGATCTTAGTCTTTGGCCAGGCGTGGTGCTCACATCT 485

QY 272 CTAATCCCGACCTTTGGGAAGCTGAGCGAGGAGTACTCTGAGGCCAGGAGTTTGAGA 331
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 484 ATAATCTAGCACTTTGGGAGGCTGAGCGAGGAGGATCACTTGAATCCAGGAGTTTGAGA 425

QY 332 CCAGGCTGGGCAATACAGGAGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 370
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 424 CCAGTCTGGGCAACACAGAGATCTCTCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 386

RESULT 8
US-09-949-016-143893/c
; Sequence 143893, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 143893
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-143893

Query Match      23.9%; Score 88.6; DB 3; Length 601;
Best Local Similarity 72.3%; Pred. No. 4.9e-20;
Matches 115; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 212 TCAGCATATGATCTCTCATATGAACCTCTCTCTGCTAGGAGTGCGAGCCTACATCG 271
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 512 TCTCCTTTTCTCCACTTCTGACAGATCTTAGTCTTTGGCCAGGCGTGGTGCTCACATCT 453

QY 272 CTAATCCCGACCTTTGGGAAGCTGAGCGAGGAGTACTCTGAGGCCAGGAGTTTGAGA 331
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 452 ATAATCTAGCACTTTGGGAGGCTGAGCGAGGAGGATCACTTGAATCCAGGAGTTTGAGA 393

QY 332 CCAGGCTGGGCAATACAGGAGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 370
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 392 CCAGTCTGGGCAACACAGAGATCTCTCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 354

RESULT 9
US-09-949-016-143894/c
; Sequence 143894, Application US/09949016
```







Db 23428 TCAAAATTATGCTTCTCGCTGGGCGTGGTGGCTCATACTTCTTATCCAGCACTTTGGGAG 23487  
Qy 293 GCTGAGGCAGGAGATTACTGGAGGCCAGGAGTTTGAGACCGGCTGGGCAATACAGCGA 352  
Db 23488 GCGAGGCAGGAGATCACTTGAGGCTAGGAGTTTGAGACCGGCTGGGCAACACAGCAA 23547  
Qy 353 GACTCTCTCTC 363  
Db 23548 GACCTATCTC 23558

Search completed: December 6, 2005, 11:47:01  
Job time : 164.493 secs

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	368.4	99.6	202381	8	AC091167	Human sapi
2	95.4	25.8	116061	8	AL732324	Human DNA
3	94.4	25.5	163863	8	AC096971	Human sapi
4	93.4	25.2	933	8	HSMTG91E10	AF076458 Homo sapi
5	93.4	25.2	81914	8	HSJD63W2	AL121906 Human DNA
C	93	25.1	56912	8	AL160176	Human DNA
	92.2	24.9	153922	8	AC011921	AL160176 Human DNA
7	92.2	24.9	174711	14	AC025515	AC011921 Homo sapi
8	92.2	24.8	99163	8	AL355602	AC025515 Homo sapi
9	91.8	24.8	165079	14	AC024517	AL355602 Human DNA
10	91.8	24.8	167300	14	AC021394	AC024517 Homo sapi
11	91.8	24.7	182509	8	AC112211	AC021394 Homo sapi
12	91.4	24.7	182509	8	AC112211	AC112211 Homo sapi
C	91.2	24.6	101769	14	AC093203	AC093203 Homo sapi
13	91.2	24.6	128540	14	AC025473	AC093203 Homo sapi
15	91.2	24.6	141633	8	AC073542	AC025473 Homo sapi
16	91.2	24.6	178042	14	AC036168	AC073542 Homo sapi
17	91.2	24.6	182892	8	AC034244	AC036168 Homo sapi
18	91.2	24.6	195068	8	AC012100	AC034244 Homo sapi
C	91.2	24.6	195068	8	AC012100	AC012100 Homo sapi

## REFERENCE AUTHORS

3 (bases 1 to 202381)  
 Birren, B., Nusbaum, C., Lander, B., Ali, A., Allen, N., Anderson, S.,  
 Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhalter, B.,  
 Camarata, J., Chang, J., Chazaro, B., Choepel, J., Collymore, A.,  
 Cooke, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J., S. Dodge, S.,  
 Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J.,  
 Gardyna, S., Gird, S., Graham, L., Grand-Pierre, N., Hafez, N.,  
 Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,  
 Kamat, A., Kartas, A., Kells, C., Landers, T., Levine, R.,  
 Lindblad-Toh, K., Liu, G., MacLean, C., Macdonald, P., Major, J.,  
 Matthews, C., McCarthy, M., Meldrum, J., Meneus, L., Mihova, T.,  
 Mengwa, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,  
 Norman, C. H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J.,  
 Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R.,  
 Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schubback, R.,  
 Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N.,  
 Stojanovic, N., Talamas, W., Tesfaye, S., Theodore, J., Tophann, K.,  
 Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X.,  
 Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.  
 Direct Submission  
 Submitted (08-JAN-2003) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 4 (bases 1 to 202381)

**TITLE**  
**JOURNAL**

Diemand, J.R., Young, S., Zainoun, P., Demkowicz, L., Zimmer, A. and Zody, M.  
Direct Submission  
Submitted (10-JAN-2003) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Jan 10, 2003 this sequence version replaced gi:27545131.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)

**COMMENT**

Research, 320 Charles Street, Cambridge, MA 02131, USA  
On Jan 10, 2003 this sequence version replaced gi:27545131.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green. P. (1996-1997)

## FEATURES

```

Location/Qualifiers
1. .202381
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="15"
/map="15"
/clone="RP11-697E2"
/clone_lib="RPCI-11 Human Male BAC"
-region
/rpt_family="MER104"
440. .729
-region

```

	/rpt_family="AlusX" complement(730..754) /rpt_family="MER104"
repeat_region	748...754 /note="<30 qual SNGL region"
unsure	793...798 /note="<30 qual SNGL region"
repeat_region	complement(801..1132) /rpt_family="AluJo"
unsure	846...851 /note="<30 qual SNGL region"
repeat_region	952...956 /note="<30 qual SNGL region"
repeat_region	1815...1860 /rpt_family="LTR16A1"
repeat_region	2145...2272 /rpt_family="LMC4a"
repeat_region	2273...2457 /rpt_family="FAM"
repeat_region	2458...2555 /rpt_family="LMC4a"
repeat_region	2556...2869 /rpt_family="Alusq"
repeat_region	2870...2883 /rpt_family="LMC4a"
repeat_region	complement(3032..3342) /rpt_family="AlusX"
repeat_region	complement(3360..3523) /rpt_family="LMA2"
repeat_region	complement(3524..3661) /rpt_family="FLAM_C"
repeat_region	complement(3662..3712) /rpt_family="LMA2"
repeat_region	3738...3917 /rpt_family="MER4E"
repeat_region	3918...4125 /rpt_family="AlusX"
repeat_region	4126...4647 /rpt_family="MER4E"
repeat_region	complement(4695..4831) /rpt_family="MR"
repeat_region	4832...4910 /rpt_family="MER81"
unsure	complement(5464..5509) /note="<30 qual SNGL region"
repeat_region	5885...6204 /rpt_family="AlusX"
repeat_region	complement(6465..6586) /rpt_family="FRAM"
repeat_region	complement(6661..7012) /rpt_family="Alusp"
repeat_region	7030...7268 /rpt_family="MER102a"
repeat_region	7270...7291 /rpt_family="(T)n"
repeat_region	complement(7322..7451) /rpt_family="FLAM_A"
unsure	complement(7407..7411) /note="<30 qual SNGL region"
repeat_region	complement(7455..7767) /rpt_family="AluJb"
unsure	complement(7512..7518) /note="<30 qual SNGL region"
unsure	complement(7690..7694) /note="<30 qual SNGL region"
unsure	complement(7803..7808) /note="<30 qual SNGL region"
unsure	complement(7824..7829) /note="<30 qual SNGL region"
repeat_region	complement(7838..8135) /rpt_family="Alusq"
repeat_region	complement(8161..8453) /rpt_family="MER82"

```

repeat_region complement(8454..8748)
repeat_region /rpt family="AluX"
repeat_region complement(8749..8921)
repeat_region /rpt family="MER82"
repeat_region 9286..9576
repeat_region /rpt family="AluJo"
repeat_region 9586..9884
repeat_region /rpt family="AluX"
repeat_region 9971..10186
repeat_region /rpt family="L1PA3"
repeat_region complement(10138..10482)
repeat_region /rpt family="MER82"
repeat_region complement(10581..10785)
repeat_region /rpt family="L1ME"
repeat_region complement(10790..11028)

Query Match 99.6%; Score 368.4; DB 8; Length 202381;
Best Local Similarity 99.7%; Pred. No. 3.8e-108;
Matches 369; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CCGTTTCAGCCACAGCAATGTATAGAGTGCTTACAACTCTTATACAGGGATACAGCAATG 60
Db 90994 CCGTTTCAGCCACAGCAATGTATAGAGTGCTTACAACTCTTATACAGGGATACAGCAATG 91053

Qy 61 TGAACACAGAGGTATGATCTGAGTTTCATCAAACTGGAAGCCATGGTAGTTATT 120
Db 91054 TGAACACAGAGGTATGATCTGAGTTTCATCAAACTGGAAGCCATGGTAGTTATT 91113

Qy 121 GTCTTTGATAGGTATGTCAACTCAACACCAATGCTCTTCTGCTGATGAGAAATGGG 180
Db 91114 GTCTTTGATAGGTATGTCAACTCAACACCAATGCTCTTCTGCTGATGAGAAATGGG 91173

Qy 181 CTGCCGATTCACAGACTTCTTGATCTTCTCAGCATATGATATCTCTCATATCAAACTC 240
Db 91174 CTGCCGATTCACAGACTTCTTGATCTTCTCAGCATATGATATCTCTCATATCAAACTC 91233

Qy 241 TCCTCTGGCTAGGAGTGCGAGCTTACATCTCTATCCAGACCTTGGGAAGCTGAGGC 300
Db 91234 TCCTCTGGCTAGGAGTGCGAGCTTACATCTCTATCCAGACCTTGGGAAGCTGAGGC 91293

Qy 301 AGGAGGATTCAGGGCCAGAGTTTGAGACCCAGGCTGGCAATACAGCGAGCTCTCT 360
Db 91294 AGGAGGATTCAGGGCCAGAGTTTGAGACCCAGGCTGGCAATACAGCGAGCTCTCT 91353

Qy 361 CTCCTAAAAA 370
Db 91354 CTCCTAAAAA 91363

```

```

RESULT 2
AL732324 116061 bp DNA linear PRI 18-MAY-2005
LOCUS Human DNA sequence from clone RP13-470G3 on chromosome X Contains
DEFINITION the 5' end of the STAR8 gene for START domain containing 8, a
ribosomal protein L31 (RPL31) pseudogene and a cytochrome c oxidase
subunit Vtc (COX6C) pseudogene, complete sequence.
ACCESSION AL732324
VERSION AL732324.6 GI:21667382
KEYWORDS HTG; COX6C; RPL31; STAR8.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 116061)
Whitehead,S.
Direct Submission
Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegasanger.ac.uk
Clone requests: clonerequest@sanger.ac.uk
On Jul 2, 2002 this sequence version replaced gi:2151526.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:

```

En., EMBL, Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at <http://www.sanger.ac.uk/Projects/C.elegans/wormpep> This sequence was generated from part of bacterial clone contigs of human chromosome X, constructed by the Sanger Centre Chromosome X Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/ChrX> RP13-470G3 is from the library RPI-13.2 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm> VECTOR: pBACE3 6

----- Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: <http://www.sanger.ac.uk>  
Contact: [vegas@sanger.ac.uk](mailto:vegas@sanger.ac.uk)

-----

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

FEATURES	Location/Qualifiers
source	1..116061 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /clone="RP13-470G3" /clone_lib="RPI-13.2" 37720..38094
gene	/locus_tag="RP13-470G3.1-001" /pseudo 37720..38094
CDS	/locus_tag="RP13-470G3.1-001" /note="match: proteins: AAH50113 CAA48925 O18602 O65071 P12947 Q9Y0T7 Q9BV33 Q9CY93 Q9GN74 Q9GP16 Q9IA76 Q9SL7 Q9U332 Q9V597 Q9XGL4" /pseudo /codon_start=1 /product="ribosomal protein L31 (RPL31) pseudogene" complement(82825..82830) complement(82947..83148) /locus_tag="RP13-470G3.2-001" /pseudo complement(82947..83148) /locus_tag="RP13-470G3.2-001" /note="match: proteins: P04038 P11950 P11951 Q9CPQ1" /pseudo /codon_start=1 /product="cytochrome c oxidase subunit Vtc (COX6C) pseudogene" join(85287..85548,93740..94172) /gene="STAR8" /locus_tag="RP13-469L5.1-002" join(85287..85548,93740..94172) /gene="STAR8" /locus_tag="RP13-469L5.1-002" /product="STAR domain containing 8" /note="match: ESTs: B114834.1"
polyA_signal	join(85290..85548,103120..103153,AL360076.9:36474..36545,AL360076.9:39928..39991,AL360076.9:40772..42189,AL360076.9:42785..42958,AL360076.9:43824..43983,AL360076.9:44484..44682,AL360076.9:45095..45306,AL360076.9:45607..45721,AL360076.9:45982..46206,AL360076.9:47186..47403,AL360076.9:47505..47681,AL360076.9:47865..49402) /gene="STAR8"
gene	/locus_tag="RP13-469L5.1-001" join(85290..85548,103120..103153,AL360076.9:36474..36545,AL360076.9:39928..39991,AL360076.9:40772..42189,AL360076.9:42785..42958,AL360076.9:43824..43983,AL360076.9:44484..44682,AL360076.9:45095..45306,AL360076.9:45607..45721,AL360076.9:45982..46206,AL360076.9:47186..47403,AL360076.9:47505..47681,AL360076.9:47865..49402) /gene="STAR8"
mRNA	/locus_tag="RP13-469L5.1-001" join(85290..85548,103120..103153,AL360076.9:36474..36545,AL360076.9:39928..39991,AL360076.9:40772..42189,AL360076.9:42785..42958,AL360076.9:43824..43983,AL360076.9:44484..44682,AL360076.9:45095..45306,AL360076.9:45607..45721,AL360076.9:45982..46206,AL360076.9:47186..47403,AL360076.9:47505..47681,AL360076.9:47865..49402) /gene="STAR8"

```
AL360076.9:44484..44682,AL360076.9:45096..45306,
AL360076.9:45607..45721,AL360076.9:45982..46206,
AL360076.9:47186..47403,AL360076.9:47505..47681,
AL360076.9:47865..49402)
/genes="STARD8"
/locus_tag="RP13-469L5.1-001"
/product="START domain containing 8"
/notes="match: ESTs: AA027102.1 AA027138.1 AA557791.1
AI752271.1 AI752272.1 AI880296.1 AW973960.1 BF828654.1
BG338127.1 BM463579.1 BM542185.1 BQ019015.1 BQ326622.1
BQ344872.1 BQ732369.1 BX102921.1 BX327875.1 CA390286.1
match: cDNAs: BC023434.1 BC034186.1 BC035587.1"
ORIGIN
Query Match      25.8%; Score 95.4; DB 8; Length 116061;
Best Local Similarity 77.6%; Pred. No. 2.1e-19;
Matches 128; Conservative 0; Mismatches 36; Indels 1; Gaps 1;
QY 12 CAGCAATGATAGAGTGCTACACCTCTATACAGGATACA-GCAATGTGAACGAGA 70
DB 83049 CAGAGCAATAGATAGAGTGCTACACCTCTACAGGAGCAAAATGCTCCAATAATGAA 83108
QY 71 GGTCAATGATAGAGTGCTTTCATCAAAACACTGGAAGCCATGGTAGTATTGCTCTTGATA 130
DB 83109 ACCATGTGCTGAGGTTTCATCAAGCACTGGAGCCATGGTAGTATTGCTCTTGATA 83168
QY 131 GGTATGTCAACCTCAACACCAATGCTCTTCTGCTGGTATGGAGAA 175
DB 83169 GGTATGTCAACCTCAACACCAATGCTCTTCTGCTGGTATGGAGAA 83213

RESULT 3
AC096971
LOCUS AC096971 163863 bp DNA linear PRI 01-AUG-2002
DEFINITION Homo sapiens chromosome 3 clone RP11-430J3, complete sequence.
ACCESSION AC096971 AC074385
VERSION AC096971.2 GI:22038614
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
REFERENCE 1 (bases 1 to 163863)
AUTHORS Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z.,
Saenphimmachak,C., Phelps,K.A., Buckley,D., Kibukawa,M., Raymond,C.
and Haugen,E.D.
Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 163863)
AUTHORS Kaul,R.K., Olson,M.V., Raymond,C., Clendenning,J. and Haugen,E.D.
TITLE Direct Submission
JOURNAL Submitted (04-OCT-2001) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
REFERENCE 3 (bases 1 to 163863)
AUTHORS Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z.,
Saenphimmachak,C., Phelps,K.A., Buckley,D., Kibukawa,M., Raymond,C.
and Haugen,E.D.
Direct Submission
JOURNAL Submitted (01-AUG-2002) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
COMMENT On Aug 1, 2002 this sequence version replaced gi:15920109.
----- Genome Center
Center: University of Washington Genome Center
Center Code: UWGC
Web site: http://www.genome.washington.edu
Contact: uwgchgs@u.washington.edu
Drafting Center: WUGSC
----- Project Information
Center project name: chr-3
Center clone name: RP11-430J3 (bc0418)
----- Summary Statistics
Sequencing vector: unknown; 51% of reads
```

```
Sequencing vector: plaamid; 49% of reads
Chemistry: Dye-terminator ET; 52% of reads
Chemistry: Dye-terminator Big Dye; 48% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 163625 bases at least Q40
Consensus quality: 163841 bases at least Q30
Consensus quality: 163861 bases at least Q20
Insert size: 163863; sum-of-contigs
Quality coverage: 7.7x in Q20 bases; sum-of-contigs
-----
Overlapping Sequences:
5': RP11-966D11 (UWGC:bc0583) AC012019
3': Mapping in progress
```

## Sequence Quality Assessment:

This entry has been annotated with sequence quality estimates computed by the Phrap assembly program. All manually edited bases have been reduced to quality zero. Quality levels above 40 are expected to have less than 1 error in 10,000 bp. Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., Phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.

## Sequence Validation:

This sequence has been validated by Multiple Complete Digest fingerprinting. Comparison of the experimentally derived digest fragments with sequence-predicted fragments is given below. The electronically-digested sequence consists of both insert and vector, in order to accurately represent the entire circular BAC. Small fragments below a variable cutoff (approximately 400-800 bp) are not resolved in the fingerprint and hence do not appear in the table. There are no significant remaining discrepancies between the experimental and predicted values. Uniquely ordered fragments are separated by dashed lines.

NsiI				EcoRI				HindIII			
SeqDerMap	FngPrnt	SeqDerMap	FngPrnt	SeqDerMap	FngPrnt	SeqDerMap	FngPrnt	SeqDerMap	FngPrnt	SeqDerMap	FngPrnt
18776	18829	8696	8829	3736	3751	---	---	---	---	---	---
2140	2139	6	<800	6382	6529	---	---	---	---	---	---
4437	4431	5304	5362	512	<800	---	---	---	---	---	---
4204	4215	971	974	449	<800	---	---	---	---	---	---
2061	2139	11515	11556	1345	1325	---	---	---	---	---	---
914	947	3018	3051	415	<800	---	---	---	---	---	---
607	<800	94	<800	1349	1325	---	---	---	---	---	---
692	<800	1496	1465	679	<800	---	---	---	---	---	---
1212	1135	5802	5835	9391	9470	---	---	---	---	---	---
2611	2648	836	822	1718	1698	---	---	---	---	---	---
2349	2351	740	822	3453	3468	---	---	---	---	---	---
1126	1135	1052	1060	706	<800	---	---	---	---	---	---



```
ORIGIN
Query Match      25.2%; Score 93.4; DB 8; Length 933;
Best Local Similarity 80.7%; Pred. No. 6.9e-19;
Matches 109; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 50 ATACGACGAATGTGAACACGAGGTCATGTCATCTGAGGTTTCATCAAAACACTGGAAGCCA 109
    |||
Db 115 ATAGTAGACTGTGGCCAGAGGCCATGATCTGAGGTTTGTCCAAACACTGGAAGCCC 174
    |||

QY 110 TGGTAGTATTGTCCTTGTATGATGATGATCAACCTCAACCAATGTCCTTCCTGGCTGAT 169
    |||
Db 175 TGGTAGTATTGTCCTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 234

QY 170 GGAGAAATGGCGTGC 184
    |||
Db 235 GGAGAAACAGACAGC 249

RESULT 5
HSDJ63M2
LOCUS
DEFINITION
HSDJ63M2 Human DNA sequence from clone RP1-63M2 on chromosome 20 Contains
the 3' end of the CBFA2T2 gene for runt domain core-binding factor
alpha subunit 2, the C20orf134 gene, the E2F1 gene for E2F
transcription factor 1, the APB2BP gene for amyloid beta (A4)
precursor protein-binding family A, member 2 protein, the C20orf144
gene and six Cpg islands, complete sequence.
AL121906
ACCESSION
VERSION AL121906.18 GI:8574371
KEYWORDS HTG; ACTIN; APB2BP; C20orf134; C20orf144; CBFA2T2; Cpg island;
E2F1; MYGRI; transcription factor.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
1 (bases 1 to 81914)
Bird.C.
Direct Submission
Submitted (13-May-2005) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk
Clone requests: clonerequests@sanger.ac.uk
On Jun 20, 2000 this sequence version replaced gi:8574056.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
En:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 20, constructed by the Sanger Centre Chromosome 20
Mapping Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr20
RP1-63M2 is from the library RPC1-1 constructed by the group of
Pleter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pCYPAC2
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: vegas@sanger.ac.uk
-----
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one subclone; and the assembly was confirmed by restriction digest,
except on the rare occasion of the clone being a YAC.
Location/Qualifiers
1. .81914
/organism="Homo sapiens"

FEATURES
Source
1. .81914
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="20"
/clone="RP1-63M2"
/clone_lib="RPC1-1"
100
misc_feature
100
gene
100
note="Clone right end: RP5-1137F22"
join(AL034421.7:88709..88750,AL034421.7:92259..92440,
210..291,1201..1454,4787..4872,6156..6351,13077..13099)
/genes="CBFA2T2"
/locus_tag="RP5-1137F22.1-006"
join(AL034421.7:88709..88750,AL034421.7:92259..92440,
210..291,1201..1454,4787..4872,6156..6351,13077..13099)
/genes="CBFA2T2"
/locus_tag="RP5-1137F22.1-006"
note="match: ESTs: BQ434235.1"
join(AL034421.7:31478..31818,AL034421.7:76100..76243,
AL034421.7:80238..80479,AL034421.7:88661..88750,
AL034421.7:92259..92440,210..391)
/genes="CBFA2T2"
/locus_tag="RP5-1137F22.1-005"
note="match: ESTs: AA635096.1 AU100534.1 BF679319.1
BM25921.1"
match: CDNAs: BC015066.1"
join(AL034421.7:76126..76243,AL034421.7:80238..80479,
AL034421.7:88661..88750,AL034421.7:92259..92440,210..369)
/genes="CBFA2T2"
/locus_tag="RP5-1137F22.1-005"
/standard_name="OTTHUMP0000030655"
/codon_start=1
/protein_id="CAI22130.1"
/db_xref="GI:56203747"
/translation="MPGSPVEVKIQSRSSPTMPPLPPINPGGPRPVSTPTALSNGL
NHSPTLNGAPSPORFSGPASSTSSALTNOQLPATCGARQLSKURFLTTLQQFGN
DISPRIGEKVRLVLALVNSTVTEIEFHCKLOEANTFPLRPFVIFPLKLANPLQREL
LHCARAAKQTPSQYLAQHEHLLNTSIASPADSSELLMEVHGNGKFPSPESCIIHLKS
MGVASRHEYFGTLOMPAHPVRLNSTLENLWDCSRSLRSTVLPFP"
join(AL034421.7:43164..43405,AL034421.7:76100..76243,
AL034421.7:80238..80479,AL034421.7:88661..88750,
AL034421.7:92259..92440,1201..1454,4787..4872,6156..6351,
13077..13145,16778..16968,20784..26473)
/genes="CBFA2T2"
/locus_tag="RP5-1137F22.1-003"
join(AL034421.7:31524..31818,AL034421.7:43118..43405,
AL034421.7:70974..71049,AL034421.7:76100..76243,
AL034421.7:80238..80479,AL034421.7:88661..88750,
AL034421.7:92259..92440,1201..1454,4787..4872,6156..6351,
13077..13145,16778..16968,20784..26473)
/genes="CBFA2T2"
/locus_tag="RP5-1137F22.1-004"
join(AL034421.7:31524..31818,AL034421.7:43118..43405,
AL034421.7:70974..71049,AL034421.7:76100..76243,
AL034421.7:80238..80479,AL034421.7:88661..88750,
AL034421.7:92259..92440,1201..1454,4787..4872,6156..6351,
13077..13145,16778..16968,20784..26473)
/genes="CBFA2T2"
/locus_tag="RP5-1137F22.1-003"
note="match: ESTs: AA459492.1 AU017082.1 AW060163.1
AW952627.1 BE741379.1 BE743916.1 BQ003927.1
BU737758.1 CA307223.1"
match: CDNAs: AF069747.1"
join(AL034421.7:43164..43405,AL034421.7:76100..76243,
AL034421.7:80238..80479,AL034421.7:88661..88750,
AL034421.7:92259..92440,1201..1454,4787..4872,6156..6351,
13077..13145,16778..16968,20784..26473)
/genes="CBFA2T2"
/locus_tag="RP5-1137F22.1-003"
note="match: ESTs: AJ948413.1 AW060163.1 BE741379.1
```





misc_feature	/db xref="RZPD:RCP1P704F20579"	
	/db xref="taxon:9606"	
polyA_site	/chromosome="20"	
	/clone="RP4-579F20"	
gene	/clone_lib="RPC1-4"	
	100	
misc_feature	/note="Clone right end: RP4-800J21"	
	complement (20584)	
polyA_site	/gene="HMG1L1"	
	complement (20585..21785)	
gene	/locus_tag="RP4-579F20.2-001"	
	complement (20585..21785)	
mRNA	/gene="HMG1L1"	
	/locus_tag="RP4-579F20.2-001"	
misc_feature	/note="match: ESTs: AA045606.1 AA050755.1 AA054783.1	
	AA093946.1 AA101746.1 AA118599.1 AA150131.1 AA196344.1	
polyA_site	AA206991.1 AA209397.1 AA220281.1 AA230725.1 AA220730.1	
	AA247271.1 AA275567.1 AA311991.1 AA313759.1 AA315282.1	
gene	AA316546.1 AA319424.1 AA330028.1 AA3357094.1 AA360132.1	
	AA360952.1 AA370554.1 AA385117.1 AA445091.1 AA459233.1	
polyA_site	AA507781.1 AA524370.1 AA541647.1 AA546572.1 AA557452.1	
	AA563485.1 AA576844.1 AA600212.1 AA607172.1 AA607305.1	
misc_feature	AA641615.1 AA641999.1 AA715713.1 AA745955.1 AA795527.1	
	AA797955.1 AA798375.1 AA845496.1 AA861823.1 AA876587.1	
polyA_site	AA994792.1 AA998198.1 AA104961.1 AA1041775.1 AA105974.1	
	AA1064990.1 AA1077928.1 AA1096960.1 AA127234.1 AA127978.1	
gene	AA151854.1 AA160153.1 AA1186025.1 AA190052.1 AA1280190.1	
	AA134636 AA148036.1 AA1351048.1 AA1368179.1 AA135682.1	
polyA_site	AA153705.1 AA1535744.1 AA1590841.1 AA1654646.1 AA1734238.1	
	AA1734851.1 AA1749800.1 AA1788272.1 AA1811093.1 AA1819512.1	
misc_feature	AA1829498.1 AA1889428.1 AA1927573.1 AA1935842.1 AA1956408.1	
	AA1963188.1 AA1998803 AA047978 AA048263 AA050935 AA011543	
polyA_site	AA014204 AA096039 AV129402 AV134319 AV171164 AV255394	
	AV301413 AA004848.1 AA016332.1 AA020136.1 AA023735.1	
misc_feature	AA026237.1 AA130136.1 AA142663.1 AA151090.1 AA160834.1	
	AA250057.1 AA258476.1 AA265014.1 AA270920.1 AA314143.1	
polyA_site	AA364611.1 AA364933.1 AA364962.1 AA364966.1 AA403673.1	
	AA464310.1 AA467569.1 AA475171.1 AA476025.1 AA537381.1	
gene	AA543443.1 AA544911.1 AA545583.1 AA547929.1 AA549623.1	
	AA553354.1 AA554198.1 AA555741.1 AA556397.1 AA556870.1	
misc_feature	AA57546.1 AA575412.1 AA575638.1 AA610999.1 D57151 P22618	
	H50055.1 H96457.1 M62207.1 M84830.1 T28966.1 W07682.1	
polyA_site	W3423.1 W85023.1 W85536.1	
	match: CDNA8: AB037850 AB037864 AF010465 AF033116 AF049239	
gene	AF049240 AF056322 AF099674 AF169946 AF169947 AF169948	
	AF201422 AF234161 AK001898 D10202 D63874 E03879 L07333	
polyA_site	L07334 L13804 L38477 L39018 M21683 M26110 M64986 M76574	
	M80436 M88749 S52624 U00431 U01343 U01345 U01348 U05593	
misc_feature	U10925 U21933 U88622 X12597 X12796 X63463 Y00463 Y14591	
	Y17153 Y17968 Z11997"	
polyA_site	complement (20612..20617)	
	/gene="HMG1L1"	
CDS	/locus_tag="RP4-579F20.2-001"	
	complement (21047..21682)	
polyA_site	/gene="HMG1L1"	
	/standard_name="OTTHUMP0000031372"	
misc_feature	/note="match: proteins: P07155 P09429 P10103 P12682 Q91596	
	Q9YH06"	
polyA_site	/codon_start=1	
	/protein_id="CAB92731.1"	
gene	/db_xref="GI:8217648"	
	/db_xref="GOA:Q9NQJ4"	
polyA_site	/db_xref="HSSP:1HMF"	
	/db_xref="InterPro:IPR000135"	
misc_feature	/db_xref="InterPro:IPR000910"	
	/db_xref="UniProt/TrEMBL:Q9NQJ4"	
polyA_site	/translation="MGKDPKPRGMSSVAFVQICREHKKKHSDASVNSSEFNK	
	CSRWKMTSAKRGEDMAKDKTYERQMKTYPKGETKKKFKDNPAPKPPSAF	
gene	FLFCSYHPKIGKEHPLGSLGDAKLGEMWNNTAADKQPGKEKKAKLKKRYKDA	
	AYQAKGPEAAKGVVKAESKKKKEEEDDEDEDEDEDEDEDEDDDDDD"	
misc_feature	join (complement (AL035541.15:871..972),	
	complement (56318..56872), complement (55723..55933),	
polyA_site	complement (51862..52032), complement (51413..51546),	
	complement (48369..48489), complement (47247..47396),	
gene	complement (45247..45407), complement (41261..41443),	
	complement (36091..36256), complement (29827..31356))	
polyA_site	/gene="CTCFL"	
	/locus_tag="RP4-579F20.1-001"	
misc_feature	join (complement (AL035541.15:871..972),	
	complement (56318..56872), complement (55723..55933),	
polyA_site	complement (51862..52032), complement (51413..51546),	
	complement (48369..48489), complement (47247..47396),	
gene	complement (45247..45407), complement (41261..41443),	
	complement (36091..36256), complement (29827..31356))	
polyA_site	/gene="CTCFL"	
	/locus_tag="RP4-579F20.1-001"	
misc_feature	/product="CCCTC-binding factor (zinc finger protein)-like"	
	/note="match: ESTs: AL732016 AL732017 AL732020 AL844280	
polyA_site	AL844303 AL844319 AL844411 AL844477 BG722482	
	match: CDNA8: AF133731 AF336042 AK001596 AK001744 U25435	
polyA_site	U51037 Z22605"	
	complement (29827)	
polyA_site	/gene="CTCFL"	
	/locus_tag="RP4-579F20.1-001"	
polyA_site	complement (29843..29848)	
	/gene="CTCFL"	
polyA_site	/locus_tag="RP4-579F20.1-001"	
	complement (join (31205..31356, 36091..36256, 41261..41443,	
polyA_site	45247..45407, 47247..47396, 48369..48489, 51413..51546,	
	51862..52032, 55723..55933, 56318..56860))	
polyA_site	/gene="CTCFL"	
	/locus_tag="RP4-579F20.1-001"	
polyA_site	/standard_name="OTTHUMP0000031376"	
	/note="match: proteins: O14709 O70162 O95015 O97600 P10078	
polyA_site	P28698 P49711 P56270 P56671 P97365 Q02975 Q08705 Q61164	
	Q61751 Q61898 Q9Y2F0"	
polyA_site	/codon_start=1	
	/product="CCCTC-binding factor (zinc finger protein)-like"	
polyA_site	/protein_id="CAI42956.1"	
	/db_xref="GI:57208876"	
polyA_site	/db_xref="GOA:Q5JU04"	
	/db_xref="InterPro:IPR007087"	
polyA_site	/db_xref="UniProt/TrEMBL:Q5JU04"	
	/translation="MAATISVLSEQFTIKELMLPEKLKEEKDGVCREKDRHP	
polyA_site	SLEAEVTSQAFQDQSVLREELVLAPESESKYILTLQTVHPTSAVELQDMSLLSI	
	QQQGVQVVOQPGQLLMLEGPQSLQOCVAISIQQLYSQEMEVLPQHALENV	
polyA_site	MVASEDQVLAETLTGLILEEOKNQLAERTKEQLFFVETMSGDSSEIVIT	
	VNSNVQEQDPTAGDAKAKSTNQRTKGAKGTFCDCVCMFTSGRMSFNHHM	
polyA_site	KTHTEKPHLCHLCTFTVTLRNHVNHTGTREPKNCDCNMAFTVSGELVHRHY	
	KTHTEKPHLCHLCTFTVTLRNHVNHTGTREPKNCDCNMAFTVSGELVHRHY	
polyA_site	SCEKPYECHITRFTQSTQTKHILQKHGENVKYPQCPCATIIARKSDLRVHMRL	
	HAYSAAELKRYCSAVFHERYALIQHOKTHKEKPKCKHCVACKQERHVAIHTH	
polyA_site	TGEKPTCLUSCNKCFRQKLLNHFYKHIDANFIPVYKSCGCKGKGFERNINURHSE	
	TKGSGEAKSAAKGRRTARRKQITLKEATKGKGAAGKEAAGDEAAAAEASTTK	
polyA_site	GEQFFGEMFPVACRETTARVKEVDEGVTCMLLNTMDK"	
	complement (join (51153..51546, 51862..52032))	
polyA_site	/gene="CTCFL"	
	/locus_tag="RP4-579F20.1-002"	
polyA_site	complement (join (51153..51546, 51862..52032))	
	/gene="CTCFL"	
polyA_site	/locus_tag="RP4-579F20.1-002"	
	/product="CCCTC-binding factor (zinc finger protein)-like"	
polyA_site	/note="match: ESTs: AW243377"	
	complement (51153)	
polyA_site	/gene="CTCFL"	
	/locus_tag="RP4-579F20.1-002"	
polyA_site	complement (51171..51176)	
	/gene="CTCFL"	
polyA_site	/locus_tag="RP4-579F20.1-002"	
	/note="Clone left end: RP4-718J7"	
polyA_site	56813	
	/note="Clone left end: RP4-718J7"	
polyA_site	Query Match 25.1%; Score 93; DB 8; Length 56912;	
	Best Local Similarity 74.5%; Pred. No. 1.2e-18;	
polyA_site	ORIGIN	
	Query Match 25.1%; Score 93; DB 8; Length 56912;	
	Best Local Similarity 74.5%; Pred. No. 1.2e-18;	

	Matches	117;	Conservative	0;	Mismatches	40;	Indels	0;	Gaps	0;
Qy	207	TTTCTCAGCATATGATCTCTCATATGAAACTCTCTCTCTGGCTAGGAGTGCGACGCTTA	266							
Db	48140	TTGCTTCGTAGATTTTTTTCAGGTAATTTAAACTCAACTCTCTGGCCAGCATGCTGGCTCA	48081							
Qy	267	CATCGCTAATCCAGCACCTTTGGGAAAGCTTGAGCGAGGAGATTTACTGGAGGCCAGGAGTT	326							
Db	48080	AACCTGTAACTCCAGCACCTTTGGGAGGCTTGAGCGAGGAGATCCCTTTGAGTCCAGGAGTT	48021							
Qy	327	TGAGACAGCGCTGGGCAATACAGCGAGACTCTCTCTC	363							
Db	48020	TGAGACAGCGCTGGTCAACATAGGGAGACCTGTCTC	47984							

RESULT	7
AC011921	
LOCUS	linear
DEFINITION	PRI 24-MAR-2002
ACCESSION	DNA
VERSION	153922 bp
KEYWORDS	Homo sapiens chromosome 17, clone RP11-15E18, complete sequence.
SOURCE	
ORGANISM	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
REFERENCE	
AUTHORS	

**TITLE**  
JOURNAL  
REFERENCE  
AUTHORS

Direct Submission  
Submitted (16 OCT-1999) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 153922)  
Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,  
Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L.,  
Bouckgalter, B., Brown, A., Camarato, M., Campotiano, A., Chang, J.,  
Chazaro, B., Chospel, Y., Colangelo, M., Collins, S., Collamore, A.,  
Cook, A., Cooke, P., DeArrellano, K., Dewar, K., Diaz, J.S., Dodge, S.,  
Faro, S., Ferreira, P., FittHugh, W., Gage, D., Galagan, J., Gardyna, S.,  
Ginde, S., Gord, S., Govette, M., Graham, L., Grand-Pierre, N.,  
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,  
Kamat, A., Karatas, A., Kells, C., Larocque, K., Lamazares, R.,  
Landers, T., Lehoczyk, J., Levine, R., Lindblad-Toh, K., Liu, G.,  
Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,  
McCarthy, M., McEwan, P., McKernan, K., Meldrim, J., Meneus, L.,  
Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,  
Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D.,  
Oliver, J., Peterson, K., Phunhkhong, P., Pierre, N., Pollara, V.,  
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,  
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R.,  
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,  
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,  
Topham, K., Traversi, M., Travis, N., Trigliio, J., Vassiliev, H.,  
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.-J., Young, G.,  
Zainoun, J., Zemlek, L., Zimmer, A. and Zody, M.  
Direct Submission

**TITLE**

**TITLE**

JOURNAL Submitted (24-MAR-2002) Whitehead Institute/MIT Center for Genomics Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT On Mar 24, 2002 this sequence version replaced gi:17977432.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>  
----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: <http://www-seq.wi.mit.edu>  
Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
----- Project Information  
Center project name: L3450  
Center clone name: 15 E 18

FEATURES	source	Location/Qualifiers
	1. .153922	
	/organism="Homo sapiens"	
	/mol_type="genomic DNA"	
	/db_xref="taxon:9606"	
	/chromosome="17"	
	/map="17"	
	/clone="RP11-15E18"	
	/clone_lib="RPC1-11 Human Male BAC"	
repeat_region	630. .938	
repeat_region	/rpt_family="AluSx"	
repeat_region	963. .1713	
repeat_region	/rpt_family="L1LMC4a"	
repeat_region	1711. .1778	
repeat_region	/rpt_family="L1M4"	
repeat_region	1779. .2089	
repeat_region	/rpt_family="AluSx"	
repeat_region	2030. .2274	
repeat_region	/rpt_family="L1M4"	
repeat_region	complement(2275. .2565)	
repeat_region	/rpt_family="AluSx"	
repeat_region	2566. .2622	
repeat_region	/rpt_family="L1M4"	
repeat_region	2623. .2926	
repeat_region	/rpt_family="AluSx"	
repeat_region	2927. .2955	
repeat_region	/rpt_family="L1M4"	
repeat_region	2956. .3270	
repeat_region	/rpt_family="AluOb"	
repeat_region	3271. .3466	
repeat_region	/rpt_family="L1M4"	
repeat_region	3899. .3931	
repeat_region	/rpt_family="AT_rich"	
repeat_region	complement(3935. .4247)	
repeat_region	/rpt_family="AluSx"	
repeat_region	5108. .5408	
repeat_region	/rpt_family="AluOb"	
repeat_region	complement(5484. .5541)	
repeat_region	/rpt_family="MIR"	
repeat_region	5616. .5637	
repeat_region	/rpt_family="AT_rich"	
repeat_region	complement(5670. .5953)	
repeat_region	/rpt_family="AluSx"	
repeat_region	complement(6006. .6304)	
repeat_region	/rpt_family="AluOb"	
repeat_region	6315. .6338	
repeat_region	/rpt_family="AT_rich"	
repeat_region	complement(6347. .6631)	
repeat_region	/rpt_family="AluOb"	
repeat_region	7751. .7852	
repeat_region	/rpt_family="HALib"	
repeat_region	complement(8253. .8546)	
repeat_region	/rpt_family="AluSx"	
repeat_region	9020. .9322	
repeat_region	/rpt_family="L2"	
repeat_region	9323. .9614	
repeat_region	/rpt_family="AluSx"	
repeat_region	9615. .9726	

repeat_region	/rpt_family="L2" complement(9798..9883)	Best Local Similarity 70.1%; Pred. No. 2.4e-18; Matches 124; Conservative 0; Mismatches 53; Indels 0; Gaps 0;
repeat_region	/rpt_family="L1MED" complement(9884..10198)	
repeat_region	/rpt_family="AluJo" complement(10269..10604)	QY 194 AGACTTCTTTGATCTTTCTCAGCATATGATCTCTCATATGAACCTCTCTCTCTGGCTAG 253       Db 90457 AGGCTGTTCCCTTGTTCTTTCATCTTTTCTCCACTAACAAAAACACCTCCCTGCTGG 90516 
repeat_region	10821..10933 /rpt_family="AluSx" complement(10982..11160)	QY 254 GAGTGGCAGCCTACATCGCTATATCCAGCACCTTGGGAAGCTGGAGCAGGAGGATTACTG 313 
repeat_region	/rpt_family="MER58B" complement(11240..11539)	Db 90517 GTGTGTGTGCTGACATCTGTAAATCCAGCATTTTGGGAGGCTTGAGCAGGAGATCACTT 90576 
repeat_region	/rpt_family="AluSc" complement(11558..11797)	QY 314 GAGGCCAGAGTTTGAGACGAGCTGGCAATACAGCAGACTCTCTCTCTAAAAA 370 
repeat_region	/rpt_family="L1MED" complement(11959..12260)	Db 90577 GAGCCCAAGATTAGAGACGAGCTGGCAACATAGACAACAATCTCTACAAA 90633 
repeat_region	/rpt_family="AluSx" complement(12340..12562)	
repeat_region	/rpt_family="L1MED" complement(12672..12942)	RESULT 8 AC025515 174711 bp DNA linear HTG 24-MAY-2002
repeat_region	13468..13489 /rpt_family="AluJo" complement(13468..13489)	LOCUS AC025515 DEFINITION Homo sapiens chromosome 17 clone RP11-634F5 map 17, WORKING DRAFT SEQUENCE, 17 unordered pieces.
repeat_region	13561..13854 /rpt_family="TAA)n" complement(14966..15038)	AC025515 4 GI:21166338 VERSION HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
repeat_region	/rpt_family="FLAM_A" complement(15138..15442)	KEYWORDS Homo sapiens (human)
repeat_region	/rpt_family="AluY" complement(15480..15516)	SOURCE Homo sapiens
repeat_region	/rpt_family="AT_rich" complement(15963..16931)	ORGANISM Homo sapiens
repeat_region	/rpt_family="MIR" complement(16998..17171)	REFERENCE 1 (bases 1 to 174711) AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
repeat_region	17235..17522 /rpt_family="AluJb" complement(17526..17841)	TITLE Homo sapiens chromosome 17, clone RP11-634F5 JOURNAL Unpublished
repeat_region	/rpt_family="MLTII" complement(18056..18175)	REFERENCE 2 (bases 1 to 174711) AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N., Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F., Boguslavskiy,L., Boukhgalter,B., Brown,A., Burkett,G., Campiolo,A., Castle,A., Choepe,Y., Colangelo,M., Collins,S., Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L., Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., LaRoque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R., Meldrim,J., Menues,L., Mihova,T., Miranda,C., Mienga,V., Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Tessaye,S., Theodore,J., Tirrell,A., Travers,M., Triggilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and Zody,W.
repeat_region	/rpt_family="FLAM_C" complement(18309..18601)	Direct Submission
repeat_region	19157..19177 /rpt_family="AT_rich" complement(20009..20496)	TITLE Submitted (09-MAR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
repeat_region	19618..19641 /rpt_family="TTTA)n" complement(20497..20686)	REFERENCE 3 (bases 1 to 174711) AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B., Brown,A., Camarata,J., Campiolo,A., Chang,J., Chazaro,B., Choepe,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzGerald,M., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRoque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G., Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Menues,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P.,
repeat_region	20497..20686 /rpt_family="MER3" complement(20687..20787)	
repeat_region	/rpt_family="L2" complement(21542..21813)	
repeat_region	/rpt_family="AluSc" complement(21884..22727)	
repeat_region	/rpt_family="L1MB4" complement(22728..23034)	
repeat_region	/rpt_family="AluJo" complement(23035..23066)	
repeat_region	/rpt_family="L1MB4" complement(23067..23261)	
repeat_region	/rpt_family="AluSx" complement(23263..23565)	
repeat_region	/rpt_family="AluSx" complement(23566..23730)	
repeat_region	/rpt_family="L1MB4" complement(23738..24312)	
repeat_region	/rpt_family="L1MB4"	
Query Match	24.98; Score 92.2; DB 8; Length 153922;	

O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigglio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.  
 Direct Submission  
 Submitted (24-MAY-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On May 24, 2002 this sequence version replaced gi:14210624.  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

# TITLE JOURNAL

## COMMENT

Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIBR  
 Web site: <http://www-seq.wi.mit.edu>  
 Contact: [sequence\\_submission@genome.wi.mit.edu](mailto:sequence_submission@genome.wi.mit.edu)  
 ----- Project Information  
 Center project name: L8265  
 Center clone name: 634\_P5

----- Summary Statistics  
 Sequencing vector: M13; M77815; 39% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.960731  
 Consensus quality: 168986 bases at least Q40  
 Consensus quality: 171447 bases at least Q30  
 Consensus quality: 172390 bases at least Q20  
 Insert size: 182000; agarose-fp  
 Insert size: 173111; sum-of-contigs  
 Quality coverage: 11.1 in Q20 bases; agarose-fp  
 Quality coverage: 11.7 in Q20.

\* NOTE: This is a 'working draft' sequence. It currently consists of 17 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

\* 1 13447: contig of 13447 bp in length  
 \* 13448 13547: gap of 100 bp  
 \* 13548 15569: contig of 2022 bp in length  
 \* 15570 15669: gap of 100 bp  
 \* 17887 17886: contig of 2217 bp in length  
 \* 17987 17986: gap of 100 bp  
 \* 20138 20137: contig of 2151 bp in length  
 \* 20238 24761: contig of 4524 bp in length  
 \* 24762 24861: gap of 100 bp  
 \* 24862 33736: contig of 8875 bp in length  
 \* 33737 33836: gap of 100 bp  
 \* 33837 40733: contig of 6897 bp in length  
 \* 40734 40833: gap of 100 bp  
 \* 40834 68844: contig of 28011 bp in length  
 \* 68845 68944: gap of 100 bp  
 \* 68945 79259: contig of 10315 bp in length  
 \* 79260 79359: gap of 100 bp  
 \* 79360 93275: contig of 13916 bp in length  
 \* 93276 93375: gap of 100 bp  
 \* 93376 100338: contig of 6963 bp in length  
 \* 100339 100438: gap of 100 bp  
 \* 100439 106496: contig of 6058 bp in length  
 \* 106497 106596: gap of 100 bp  
 \* 106597 114601: contig of 8005 bp in length  
 \* 114602 114701: gap of 100 bp  
 \* 114702 135193: contig of 20492 bp in length  
 \* 135194 135293: gap of 100 bp  
 \* 135294 153360: contig of 18067 bp in length  
 \* 153361 153460: gap of 100 bp

\* 153461 170470: contig of 17010 bp in length  
 \* 170471 170570: gap of 100 bp  
 \* 170571 174711: contig of 4141 bp in length.  
 FEATURES  
 source  
 1..174711  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="17"  
 /map="17"  
 /clone="RP11-634P5"  
 /clone\_lib="RPC1-11 Human Male BAC"  
 1..13447  
 /note="assembly\_fragment"  
 /clone\_end:SP6  
 vector\_side:left  
 13448..13547  
 /estimated\_length=100  
 13548..15569  
 /note="assembly\_fragment"  
 15570..15669  
 /estimated\_length=100  
 15670..17886  
 /note="assembly\_fragment"  
 17887..17986  
 /estimated\_length=100  
 17987..20137  
 /note="assembly\_fragment"  
 20138..20237  
 /estimated\_length=100  
 20238..24761  
 /note="assembly\_fragment"  
 24762..24861  
 /estimated\_length=100  
 24862..33736  
 /note="assembly\_fragment"  
 33737..33836  
 /estimated\_length=100  
 33837..40733  
 /note="assembly\_fragment"  
 40734..40833  
 /estimated\_length=100  
 40834..68844  
 /note="assembly\_fragment"  
 68845..68944  
 /estimated\_length=100  
 68945..79259  
 /note="assembly\_fragment"  
 79260..79359  
 /estimated\_length=100  
 79360..93275  
 /note="assembly\_fragment"  
 93276..93375  
 /estimated\_length=100  
 93376..100338  
 /note="assembly\_fragment"  
 100339..100438  
 /estimated\_length=100  
 100439..106496  
 /note="assembly\_fragment"  
 106497..106596  
 /estimated\_length=100  
 106597..114601  
 /note="assembly\_fragment"  
 114602..114701  
 /estimated\_length=100  
 114702..135193  
 /note="assembly\_fragment"  
 135194..135293  
 /estimated\_length=100  
 135294..153360  
 /note="assembly\_fragment"

```
Query Match      24.9%; Score 92.2; DB 14; Length 174711;
Best Local Similarity 70.1%; Pred. No. 2.4e-18;
Matches 124; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 194 AGACTTCTTGATCTTCTCCTCAGCATATGACTCTCATATGAACACTCTCTCTCTGCTAG 253
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8124 AGGCTGTCTCTTGTCTTCTCATCTTACTGTCTCCACTAACAAAAACACCTCTCTGCTGG 8183

QY 254 GAGTGGCAGCTCATCGCTTAATCCAGCAGCTTGGAGAGCTGAGCAGGAGGATTAAGT 313
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8184 GTGTGGTGGCTGACATCTGTAATCCAGCATTTTGGAGAGCTGAGCAGGAGGATCACTT 8243

QY 314 GAGCCAGGAGTTTGAGCAGCAGGCTGGGCAATACAGGAGACTCTCTCTCTTAATAA 370
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8244 GAGCCCAAGAGTTAGACACAGCTGGGCAACATAGCAACACATCTCTACAAA 8300

RESULT 9
AL355602      99163 bp      DNA      linear      PRI 18-MAY-2005
LOCUS
DEFINITION
Human DNA sequence from clone RP5-1166F10 on chromosome 1 Contains
the 3' end of a novel gene (PLJ36179), complete sequence.
ACCESSION
AL355602.11 GI:13560009
VERSION
HTG; FLJ36179.
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 99163)
Direct Submission
Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk
Clone requests: clonerequest@sanger.ac.uk
On Apr 6, 2001 this sequence version replaced gi:13396512.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr1
RP5-1166F10 is from the library RPCI-5 constructed by the group of
Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pCYPAC2

----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: vegas@sanger.ac.uk

-----
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one subclone; and the assembly was confirmed by restriction digest,
except on the rare occasion of the clone being a YAC.
Location/Qualifiers
1 . 99163
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="RZPD:RPCIF704F101166"
/db_xref="taxon:9606"
/chromosome="1"
/clones="RP5-1166F10"
/clone_lib="RPCI-5"

FEATURES
source
misc_feature
1
```

```
/note="Clone left end: RP5-1166F10"
join(AL591916.8:151507..152088,AL591916.8:155009..155043,
AL591916.8:155768..156004,5057..5484,5645..5745,
6291..6984)
/locus_tag="RP5-1166F10.1-001"
join(AL591916.8:151507..152088,AL591916.8:155009..155043,
AL591916.8:155768..156004,5057..5484,5645..5745,
6291..6984)
/locus_tag="RP5-1166F10.1-001"
/product="novel transcript"
/note="match: ESTs: Em:AW628519.1
match: cDNAs: Em:AK093498.1 Em:AX748168.1"
6956..6961
/locus_tag="RP5-1166F10.1-001"
6984
/locus_tag="RP5-1166F10.1-001"
99064
/note="Clone left end: RP1-37J18"

ORIGIN
Query Match      24.8%; Score 91.8; DB 8; Length 99163;
Best Local Similarity 77.6%; Pred. No. 3.1e-18;
Matches 111; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 228 TCATATGAAACTCTCTCTGCTAGGAGTGGCAGCCTACATCGCTAATCCAGCACCTT 287
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 12266 TCTTATAAAGAGACCCCTGCTGGCATGTGGCTCAGCTGTATCCAGCACCTTT 12325

QY 288 GGAAGCTCAGGAGGAGGATTACTGGAGCCAGGAGTTTGAGACCAGGCTGGCAATAC 347
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 12326 GGGAGCCGAGGAGGAGGATCACTGGAGCCAGGAGTTCAAGACAGCAGCTGGTAACAC 12385

QY 348 AGGAGACTCTCTCTCTTAATAA 370
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 12386 AGCAGACTCTCTCTCAAAAAA 12408

RESULT 10
AC024517      165079 bp      DNA      linear      HTG 21-MAR-2000
LOCUS
DEFINITION
Homo sapiens chromosome 1 clone RP11-168B8 map 1, WORKING DRAFT
SEQUENCE, 30 unordered pieces.
ACCESSION
AC024517.2 GI:7272182
VERSION
HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 165079)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Beda,F., Boguslavsky,L.,
Boukhgalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,
Choepe,Y., Collangelo,M., Collins,S., Collamore,A., Cooke,P.,
Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Domino,M., Doyle,M.,
Feneator,J., Ferreira,P., FitzHugh,W., Forrest,C., Gage,D.,
Galegan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., Landers,T., Largocque,K., Lehoczy,J., Levine,R.,
Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N., McCarthy,M.,
McEwan,P., McGurk,A., McKernan,K., McPheeters,R., Meldrum,J.,
Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J., Naylor,J.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Olivari,T.M.,
Peterson,K., Pierre,N., Pisani,C., Pollara,V., Raymond,C.,
Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S.,
Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A.,
```

Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.

**TITLE**  
JOURNAL  
COMMENT

Direct Submission  
Submitted (28-FEB-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Mar 21, 2000 this sequence version replaced GI:108313.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: <http://www-seq.wi.mit.edu>  
Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)

----- Project Information  
Center project name: L7252  
Center clone name: 168\_B8

----- Summary Statistics  
Sequencing vector: M13; M77815; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 144165 bases at least Q40  
Consensus quality: 153730 bases at least Q30  
Consensus quality: 158511 bases at least Q20  
Insert size: 170000; agarose-fp  
Quality coverage: 3.2 in Q20 bases; agarose-fp  
Quality coverage: 3.4 in Q20 bases; sum-of-contigs

-----  
\* NOTE: This is a 'working draft' sequence. It currently consists of 30 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

\* 1 1125: contig of 1125 bp in length  
\* 1126 1225: gap of 100 bp  
\* 1226 2573: contig of 1348 bp in length  
\* 2574 2673: gap of 100 bp  
\* 2674 3825: contig of 1152 bp in length  
\* 3826 3925: gap of 100 bp  
\* 3926 5196: contig of 1271 bp in length  
\* 5197 5296: gap of 100 bp  
\* 5297 6379: contig of 1083 bp in length  
\* 6380 6479: gap of 100 bp  
\* 6480 8499: contig of 2020 bp in length  
\* 8500 8599: gap of 100 bp  
\* 8600 11025: contig of 2426 bp in length  
\* 11026 11125: gap of 100 bp  
\* 11126 13438: contig of 2313 bp in length  
\* 13439 16281: contig of 100 bp  
\* 16282 16381: gap of 100 bp  
\* 16382 18437: contig of 2056 bp in length  
\* 18438 18537: gap of 100 bp  
\* 18538 22277: contig of 3740 bp in length  
\* 22278 22377: gap of 100 bp  
\* 22378 25101: contig of 2724 bp in length  
\* 25102 25201: gap of 100 bp  
\* 25202 27806: contig of 2605 bp in length  
\* 27807 27906: gap of 100 bp  
\* 27907 30832: contig of 2926 bp in length  
\* 30833 30933: gap of 100 bp  
\* 30934 33715: contig of 2782 bp in length  
\* 33716 33815: gap of 100 bp  
\* 33816 3617: contig of 2803 bp in length  
\* 36178 36717: gap of 100 bp  
\* 36718 39412: contig of 2695 bp in length  
\* 39413 39512: gap of 100 bp  
\* 39513 43281: contig of 3769 bp in length

\* 43282 43381: gap of 100 bp  
\* 43382 48656: contig of 5275 bp in length  
\* 48657 53648: gap of 100 bp  
\* 53649 53748: contig of 4892 bp in length  
\* 53749 57878: gap of 100 bp  
\* 57879 57978: contig of 4130 bp in length  
\* 57979 63627: gap of 100 bp  
\* 63628 63728: contig of 5649 bp in length  
\* 63729 69087: gap of 100 bp  
\* 69088 69187: contig of 5360 bp in length  
\* 69188 77490: gap of 100 bp  
\* 77491 77590: contig of 8303 bp in length  
\* 77591 85777: gap of 100 bp  
\* 85778 93590: contig of 8187 bp in length  
\* 93591 93690: gap of 100 bp  
\* 93691 102643: contig of 7713 bp in length  
\* 102644 102743: gap of 100 bp  
\* 102744 114758: contig of 12015 bp in length  
\* 114759 114858: gap of 100 bp  
\* 114859 133243: contig of 18385 bp in length  
\* 133244 133344: gap of 100 bp  
\* 133345 165073: contig of 31736 bp in length.

FEATURES  
source  
1. .165079  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="1"  
/map="11"  
/clones="RP11-168B8"  
/clone\_lib="RPC1-11 Human Male BAC"  
1. .1125  
/note="assembly\_fragment"  
1126. .1225  
/estimated\_length=100  
1226. .2573  
/note="assembly\_fragment"  
2574. .2673  
/estimated\_length=100  
2674. .3825  
/note="assembly\_fragment"  
3826. .3925  
/estimated\_length=100  
3926. .5196  
/note="assembly\_fragment"  
5197. .5296  
/estimated\_length=100  
5297. .6379  
/note="assembly\_fragment"  
6380. .6479  
/estimated\_length=100  
6480. .8499  
/note="assembly\_fragment"  
8500. .8599  
/estimated\_length=100  
8600. .11025  
/note="assembly\_fragment"  
11026. .11125  
/estimated\_length=100  
11126. .13438  
/note="assembly\_fragment"  
13439. .13538  
/estimated\_length=100  
13539. .16281  
/note="assembly\_fragment"  
clone\_end:T7  
vector\_side:left  
16282. .16381  
/estimated\_length=100  
16382. .18437  
/note="assembly\_fragment"



```

gap      18438..18537
/estimated_length=100
misc_feature
18538..22277
/notes="assembly_fragment"
gap      22278..22377
/estimated_length=100
misc_feature
22378..25101
/notes="assembly_fragment"
gap      25102..25201
/estimated_length=100
misc_feature
25202..27806
/notes="assembly_fragment"
gap      27807..27906
/estimated_length=100
misc_feature
27907..30832
/notes="assembly_fragment"
gap      30833..30932
/estimated_length=100
misc_feature
30933..33714
/notes="assembly_fragment"

Query Match      24.8%; Score 91.8; DB 14; Length 165079;
Best Local Similarity 77.6%; Pred. No. 3.3e-18;
Matches 111; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 228 TCATATGAACTCTCTCTGGCTAGGAGTGGCGAGCTACATCGCTAAATCCGAGCACCTT 287
Db 45608 TCTTATAAAGAGACCCCTGGCTGGCATGGTGGCTCACTGTAATCCGAGCACCTT 45667

QY 288 GGGAAGCTGAGGAGGAGGAGTACTGGAGGCCAGGAGTTTGAGACAGGCGTGGGCAATAC 347
Db 45668 GGAGGCCGAGGAGGAGGAGTACTGGAGGCCAGGAGTTTGAGACAGGCGTGGGCAATAC 45727

QY 348 AGCGAGACTCTCTCTCTCTAAAAA 370
Db 45728 AGCAAGACTCTCTCTCAAAAAA 45750

RESULT 11
AC021394/c
LOCUS
DEFINITION Homo sapiens clone RP11-28E15, WORKING DRAFT SEQUENCE, 11 unordered
pieces.
ACCESSION AC021394
VERSION AC021394.3 GI:7630674
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 167300)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens, clone RP11-28E15
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 167300)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
Boguelavkiy,L., Bouhgalter,B., Brown,A., Burkett,G., Castle,A.,
Choepe,Y., Collangelo,M., Collins,S., Collymore,A., Cooke,P.,
DeArellano,K., Dewar,K., Domino,M., Doyle,M., Fenesstor,J.,
Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galegan,J.,
Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G., Locke,K.,
Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
McPheeters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J.,
Norman,C.H., O'Connor,T., O'Donnell,P., Olivar,T.M., Peterson,K.,
Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
Zimmer,A. and Zody,M.

```

# TITLE JOURNAL COMMENT

Direct Submission  
Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Apr 21, 2000 this sequence version replaced gi:6850427.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html  
----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence submissions@genome.wi.mit.edu  
----- Project Information  
Center project name: L4746  
Center clone name: 28\_E\_15  
----- Summary Statistics  
Sequencing vector: M13; M77815; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 158720 bases at least Q40  
Consensus quality: 162531 bases at least Q30  
Insert size: 167000; agarose-fp  
Insert size: 166300; sum-of-contigs  
Quality coverage: 4.7 in Q20 bases; agarose-fp  
Quality coverage: 4.7 in Q20 bases; sum-of-contigs  
-----

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 11 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence.  
\* as soon as it is available and the accession number will  
\* be preserved.

\* 1 1018: contig of 1018 bp in length  
\* 1019 1118: gap of 100 bp  
\* 1119 4175: contig of 3057 bp in length  
\* 4176 4275: gap of 100 bp  
\* 4276 8989: contig of 4714 bp in length  
\* 8990 9089: gap of 100 bp  
\* 9090 14343: contig of 5254 bp in length  
\* 14344 14443: gap of 100 bp  
\* 14444 20389: contig of 5946 bp in length  
\* 20390 20489: gap of 100 bp  
\* 20490 28508: contig of 8019 bp in length  
\* 28509 28608: gap of 100 bp  
\* 28609 37887: contig of 9279 bp in length  
\* 37888 37988: gap of 100 bp  
\* 37988 54331: contig of 16344 bp in length  
\* 54332 54431: gap of 100 bp  
\* 54432 84316: contig of 29885 bp in length  
\* 84317 84416: gap of 100 bp  
\* 84417 121379: contig of 36963 bp in length  
\* 121380 121479: gap of 100 bp  
\* 121480 167300: contig of 45821 bp in length.

FEATURES  
source  
1..167300  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone="RP11-28E15"  
/clone\_lib="RPCI-11 Human Male BAC"  
misc\_feature  
1..1018  
/note="assembly\_fragment"  
gap  
1019..1118  
/estimated\_length=100  
misc\_feature  
1119..4175  
/note="assembly\_fragment"  
gap  
4176..4275  
/estimated\_length=100  
misc\_feature  
4276..8989  
/note="assembly\_fragment"



```

gap          8990..9089
misc_feature /estimated_length=100
9090..14343
/Note="assembly_fragment"
gap          14344..14443
misc_feature /estimated_length=100
14444..20389
/Note="assembly_fragment"
clone_end:T7
vector_side:left"
20390..20489
misc_feature /estimated_length=100
20490..28508
/Note="assembly_fragment"
gap          28509..28608
misc_feature /estimated_length=100
28609..37887
/Note="assembly_fragment"
clone_end:SP6
vector_side:right"
37888..37987
gap          /estimated_length=100
37988..54331
misc_feature /Note="assembly_fragment"
54332..54431
gap          /estimated_length=100
54432..84316
misc_feature /Note="assembly_fragment"
84317..84416
gap          /estimated_length=100
84417..121379
misc_feature /Note="assembly_fragment"
121380..121479
gap          /estimated_length=100
121480..167300
misc_feature /Note="assembly_fragment"

ORIGIN
Query Match      24.8%; Score 91.8; DB 14; Length 167300;
Best Local Similarity 77.6%; Pred. No. 3.3e-18;
Matches 111; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 228 TCATATGAACCTCTCTCTGCTAGGAGTGGCAGCCTACATCGCTAATCCAGCACCTT 287
Db 13204 TCTTATAAAGAGACCCCTGCTGGCATGCTGCTCACATCTGTATATCCAGCACCTT 13145

QY 288 GGGAGCTGAGCAGGAGATTACTGGAGGCCAGGAGTTTGAGACCAGGCTGGGCAATAC 347
Db 13144 GGGAGGCCGAGCAGGAGATCACTGGAGCCAGGAGTTCAAGACCAGCGTGGTAAACAC 13085

QY 348 AGCGAGACTCTCTCTCCTAATAAA 370
Db 13084 AGCAAGACTCTGCTCTCAAAAAAA 13062

RESULT 12
AC112211/c AC112211 182509 bp DNA linear PRI 24-MAY-2002
LOCUS Homo sapiens chromosome 3 clone RP11-627J17, complete sequence.
DEFINITION AC112211
ACCESSION AC112211
VERSION AC112211.2 GI:21166208
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
REFERENCE 1 (bases 1 to 182509)
AUTHORS Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z.,
Saenphimmachak,C., Phelps,K.A., Buckley,D., Raymond,C. and
Haugen,E.D.
Direct Submission
TITLE (bases 1 to 182509)
JOURNAL Submitted (20-FEB-2002) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
REFERENCE 3 (bases 1 to 182509)
AUTHORS Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z.,
Saenphimmachak,C., Phelps,K.A., Buckley,D., Raymond,C. and
Haugen,E.D.
Direct Submission
TITLE (bases 1 to 182509)
JOURNAL Submitted (24-MAY-2002) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
COMMENT On May 24, 2002 this sequence version replaced gi:18767508.
----- Genome Center
Center: University of Washington Genome Center
Center Code: UWGC
Web site: http://www.genome.washington.edu
Contact: uwgctgs@u.washington.edu
----- Project Information
Center project name: chr-3
Center clone name: RP11-627J17 (bc0502)
----- Summary Statistics
Sequencing vector: plasmid; 100% of reads
Chemistry: Dye-terminator ET; 50% of reads
Chemistry: Dye-terminator Big Dye; 50% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 182361 bases at least Q40
Consensus quality: 182497 bases at least Q30
Consensus quality: 182509 bases at least Q20
Insert size: 182509; sum-of-contigs
Quality coverage: 7.9x in Q20 bases; sum-of-contigs
-----
Overlapping Sequences:
5': Mapping in progress
3': Mapping in progress
-----
Sequence Quality Assessment:
This entry has been annotated with sequence quality
estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than
1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the
GenBank flat file format but are available as part
of this entry's ASN.1 file.
This sequence was finished as follows unless otherwise noted:
all regions were either double-stranded or sequenced with an
alternate chemistry or covered by high quality data (i.e., Phred
quality >= 30); an attempt was made to resolve all sequencing
problems, such as compressions and repeats; all regions were
covered by at least one plasmid subclone or more than one M13
subclone; and the assembly was confirmed by restriction digest.
-----
Sequence Validation:
This sequence has been validated by Multiple Complete Digest
fingerprinting. Comparison of the experimentally derived digest
fragments with sequence-predicted fragments is given below.
The electronically-digested sequence consists of both insert and
vector, in order to accurately represent the entire circular BAC.
Small fragments below a variable cutoff (approximately 400-800 bp)
are not resolved in the fingerprint and hence do not appear
in the table. There are no significant remaining discrepancies
between the experimental and predicted values. Uniquely ordered
fragments are separated by dashed lines.
-----
ECORI BglII
SeqDerMap FngPrnt SeqDerMap FngPrnt SeqDerMap FngPrnt
-----
8696 8777 4710 4671 2282 2251
-----
Unpublished

```

6	<800	6382	6514	2067	2068	1882	1842	1568	1541	2199	2251
8311	8264	512	<800	5626	5794	1860	1842	3047	3110	1365	1358
3794	3847	449	<800	4289	4117	1496	1500	1202	1181	334	<800
517	<800	1417	1461	4382	4412	131	<800	1526	1461	1620	1585
5654	5717	5366	5126	4594	4580	1709	1842	1589	1541	4068	3991
3530	3552	590	<800	10637	10624	1131	1135	1882	1928	10154	10136
2635	2737	3606	3651	4577	4580	1365	1329	2629	2732	1586	1585
3532	3552	274	<800	246	<800	25	<800	4105	4080	5883	5490
12069	11650	4570	4539	933	948	25	<800	675	<800	3347	3348
8046	8264	9049	8995	15529	15666	1046	1051	6992	6962	961	948
2974	3009	4374	4320	184	<800	340	<800	15952	16085	1952	1961
3451	3552	4978	5126	4296	4283	1313	1329	1482	1461	1371	1358
2090	2126	11396	11164	2736	2818	1294	1329	6532	6514	521	<800
1581	1500	1746	1719	3990	3991	9423	9329	8901	8995	454	<800
555	<800	1599	1541	4246	4117	281	<800	3206	3280	293	<800
210	<800	8101	8204	693	<800	6458	6434	192	<800	-----	-----
1074	1051	2394	2339	3597	3667	-----	-----	7709	7507	-----	-----
2455	2540	1944	1928	12762	12592	-----	-----	294	<800	-----	-----
8853	8777	5118	5126	1327	1358	-----	-----	-----	-----	-----	-----
361	<800	3069	3110	3651	3667	-----	-----	-----	-----	-----	-----
3186	3259	2276	2339	1831	1785	-----	-----	-----	-----	-----	-----
19747	19152	7388	7507	404	<800	-----	-----	-----	-----	-----	-----
1669	1668	196	<800	3353	3348	-----	-----	-----	-----	-----	-----
1367	1329	312	<800	159	<800	-----	-----	-----	-----	-----	-----
3124	3171	2294	2339	4303	4283	-----	-----	-----	-----	-----	-----
24350	24894	5185	5126	483	<800	-----	-----	-----	-----	-----	-----
6070	6075	714	<800	4395	4412	-----	-----	-----	-----	-----	-----
1385	1329	1355	1358	6197	6227	-----	-----	-----	-----	-----	-----
1812	1842	8240	8204	3312	3348	-----	-----	-----	-----	-----	-----
865	863	1278	1358	9314	9271	-----	-----	-----	-----	-----	-----
839	863	1678	1627	3580	3667	-----	-----	-----	-----	-----	-----
2111	2126	1968	1928	5491	5490	-----	-----	-----	-----	-----	-----
1164	1135	188	<800	5454	5490	-----	-----	-----	-----	-----	-----
4781	4748	463	<800	560	<800	-----	-----	-----	-----	-----	-----
655	<800	42	<800	6499	6511	-----	-----	-----	-----	-----	-----
6026	6075	5102	5126	4391	4412	-----	-----	-----	-----	-----	-----
1883	1842	1401	1461	3029	3066	-----	-----	-----	-----	-----	-----

FEATURES

Location/Qualifiers

Query Match 24.7%; Score 91.4; DB 8; Length 182509;  
Best Local Similarity 75.8%; Pred. No. 4.4e-18;  
Matches 113; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY

222 ATACTCTCATATCAAACTCTCTCTCTGGCTAGGAGTGCGAGCTACATCGCTAATCCCAG 281

19898 ATCCTCTCAATAATATTCAACGTCAGGCTAGGAGTGCGAGCTCAGGCTGTAAATCCCAG 19839

QY

282 CACCTTGGGAAGCTGAGGAGGAGGATTACTTGAGGCCAGGAGTTTGAGACCGAGCTGGG 341

19838 CACTTTGGGAGCTGAGGTGGGAGGATTACTTCAGCCATGGAGTTCAAGACCGAGCTGGG 19779

QY

342 CAATACAGCGAGACTCTCTCTCTCAAAAA 370

19778 CAACACAGGAGACCTGTCTCTACAAAA 19750

Db

RESULT 13

AC093203/c

LOCUS

DEFINITION

AC093203

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

AC093203

101769 bp DNA linear HTG 07-MAR-2002

Homo sapiens chromosome 5 clone CTC-262A8, WORKING DRAFT SEQUENCE,

2 ordered pieces.

AC093203

AC093203.4 GI:19224825

HTG; HTGS\_PHASE2; HTGS\_DRAFT; HTGS\_ACTIVEFIN.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 101769)

DOE Joint Genome Institute.

Sequencing of Human Chromosome 5

Unpublished

```

REFERENCE
AUTHORS      2 (bases 1 to 101769)
TITLE        DOE Joint Genome Institute.
JOURNAL      Submitted (16-AUG-2001) Production Sequencing Facility, DOE Joint
              Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE
AUTHORS      3 (bases 1 to 101769)
TITLE        DOE Joint Genome Institute.
JOURNAL      Submitted (07-MAR-2002) Production Sequencing Facility, DOE Joint
              Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
              On Mar 7, 2002 this sequence version replaced gi:15426059.
COMMENT      -----Genome Center
              Center: Joint Genome Institute
              Web site: http://www.jgi.doe.gov
              -----
              Project Information
              Center Project Name: 285795
              Center clone name: CIT-HSPC_262A8
              -----
              Summary Statistics
              Consensus quality: 101567 bases at least Q40
              Consensus quality: 101611 bases at least Q30
              Consensus quality: 101641 bases at least Q20
              Estimated insert size: 104000; agarose-fp estimation
              Estimated insert size: 34.54 in Q20 bases; agarose-fp estimation
              Quality coverage: 35.33 in Q20 bases; sum-of-contigs estimation.
              * NOTE: This is a 'working draft' sequence. It currently
              * consists of 2 contigs. Gaps between the contigs
              * are represented as runs of N. The order of the pieces
              * is believed to be correct as given, however the sizes
              * of the gaps between them are based on estimates that have
              * provided by the submittor.
              * This sequence will be replaced
              * by the finished sequence as soon as it is available and
              * the accession number will be preserved.
              *
              1 12059: contig of 12059 bp in length
              *
              12060 12159: gap of unknown length
              *
              12160 101769: contig of 89610 bp in length.
              *
              Location/Qualifiers
                1..101769
                  /organism="Homo sapiens"
                  /mol_type="genomic DNA"
                  /db_xref="taxon:9606"
                  /chromosome="5"
                  /clone="CTC-262A8"
                  /clone_lib="CalTech human BAC library C"
                  12060..12159
                    /estimated_length=unknown
              gap
              ORIGIN
              Query Match      24.6%; Score 91.2; DB 14; Length 101769;
              Best Local Similarity 71.4%; Pred. NO. 4.9e-18;
              Matches 137; Conservative 0; Mismatches 48; Indels 7; Gaps 1;
              QY      2 GGTTCAGCCACGCAAGTGTATAGTGTCTACAC-----CCTCTATACAGGGATACA 54
              DB      74874 GTTTCAGCCACGCAAGTGTATAGAGAGCTACAAATATCTCCAACTATGAAATCAC 74815
              QY      55 GCAATGTGAAACACAGAGTCTATGTCATCTGAGGTTTTCATCAAAACACTGGAAGCCATGGTA 114
              DB      74814 AGACACTGGTCAGAGGCCAAGCACTGAGGTTCTGTCAAGCTCTCGAAGCCCTGGTA 74755
              QY      115 GTTATGTCTTGTAGTAGTATGTCAACTCTCAACCAATGTCCTTCCTGGCTGATGGAGA 174
              DB      74754 GTTACTGTCTTAAATTGTATGTCTCAGTCTCAACCAATGCGCTTCCTGGCTGATGGAGA 74695
              QY      175 AATGGGCTGCCC 186
              DB      74694 AAGGGATGCCCC 74683

```

---

```

RESULT 14
AC025473/c
LOCUS
DEFINITION Homo sapiens chromosome 5 clone RP11-183N2, WORKING DRAFT SEQUENCE,
              11 unordered pieces.
AC025473
VERSION      AC025473.2 GI:7417674
KEYWORDS     HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE       Homo sapiens
ORGANISM     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
              Hominiidae; Homo.
REFERENCE
AUTHORS      1 (bases 1 to 128540)
TITLE        DOE Joint Genome Institute.
JOURNAL      Sequencing of Human Chromosome 5
REFERENCE
AUTHORS      2 (bases 1 to 128540)
TITLE        DOE Joint Genome Institute.
JOURNAL      Direct Submission
              Submitted (09-MAR-2000) Production Sequencing Facility, DOE Joint
              Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
              On Apr 5, 2000 this sequence version replaced gi:7211967.
COMMENT      -----Genome Center
              Center: Joint Genome Institute
              Web site: http://www.jgi.doe.gov
              -----
              Summary Statistics
              Consensus quality: 118729 bases at least Q40
              Consensus quality: 125307 bases at least Q30
              Consensus quality: 126460 bases at least Q20
              Estimated insert size: 128540; sum-of-contigs estimation
              Estimated insert size: 202600; agarose-fp estimation
              Quality coverage: 3.58x in Q20 bases; agarose-fp estimation
              Quality coverage: 5.64x in Q20 bases; sum-of-contigs estimation
              -----
              * NOTE: This is a 'working draft' sequence. It currently
              * consists of 11 contigs. The true order of the pieces
              * is not known and their order in this sequence record is
              * arbitrary. Gaps between the contigs are represented as
              * runs of N, but the exact sizes of the gaps are unknown.
              * This record will be updated with the finished sequence
              * as soon as it is available and the accession number will
              * be preserved.
              *
              1 2361: contig of 2361 bp in length
              *
              2362 4766: contig of 2405 bp in length
              *
              4767 7300: contig of 2534 bp in length
              *
              7301 10400: contig of 3100 bp in length
              *
              10401 14252: contig of 3852 bp in length
              *
              14253 20943: contig of 6691 bp in length
              *
              20944 26181: contig of 5238 bp in length
              *
              26182 32137: contig of 5956 bp in length
              *
              32138 47316: contig of 15179 bp in length
              *
              47317 71055: contig of 23739 bp in length
              *
              71056 128540: contig of 57485 bp in length.
              *
              Location/Qualifiers
                1..128540
                  /organism="Homo sapiens"
                  /mol_type="genomic DNA"
                  /db_xref="taxon:9606"
                  /chromosome="5"
                  /clone="RP11-183N2"
              FEATURES
              source
              ORIGIN

```

Query Match 24.6%; Score 91.2; DB 14; Length 128540;  
Best Local Similarity 71.4%; Pred. No. 5e-18; Indels 7; Gaps 1;  
Matches 137; Conservative 0; Mismatches 48; Indels 7; Gaps 1;

QY 2 GGTTCAGCCACAGCAAAATGTATAGAGTGCTACAAC-----CCTCTATACAGGGATACA 54  
DB 128064 GGTTCAGCCACAGCAAAAGTTTATAAGAGCTACAACAATACTCCAACAATATGAAATCAC 114  
QY 55 GCAATGTGAACACGAGGTCAATGATCTGAGGTTTCATCAAAACACTGGAAGCCATGGTA 114  
DB 128005 GGTTCAGCCACAGCAAAAGTTTATAAGAGCTACAACAATACTCCAACAATATGAAATCAC 114  
QY 115 GCAATGTGAACACGAGGTCAATGATCTGAGGTTTCATCAAAACACTGGAAGCCATGGTA 174  
DB 128004 AGACACTTGGTCAGAGGCCAAGCAACTGAGGTTCTGCAAGCTCTGGAAGCCCTGGTA 127945  
QY 115 GTTATTGTCCTTGTATAGGTATGCAACCTCAACCAATGTCCTTCTGGCTCATGGAGA 174  
DB 127944 GTTATTGTCCTTGTATAGGTATGCAACCTCAACCAATGTCCTTCTGGCTCATGGAGA 174  
QY 175 AATGGGCTGCCC 186  
DB 127884 AAGGATGGCCC 127873

RESULT 15  
AC073542 141633 bp DNA linear PRI 01-MAY-2001  
LOCUS Homo sapiens chromosome 5 clone RP11-221L16, complete sequence.  
DEFINITION AC073542  
ACCESSION AC073542 GI:13876495  
VERSION HTG.  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM  
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE 1 (bases 1 to 141633)  
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.  
TITLE Direct Submission  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 141633)  
AUTHORS DOE Joint Genome Institute.  
TITLE Direct Submission  
JOURNAL Submitted (22-JUN-2000) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
REFERENCE 3 (bases 1 to 141633)  
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.  
TITLE Direct Submission  
JOURNAL Submitted (28-APR-2001) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
REFERENCE 4 (bases 1 to 141633)  
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.  
TITLE Direct Submission  
JOURNAL Submitted (01-MAY-2001) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
COMMENT On Apr 28, 2001 this sequence version replaced gi:13677059.  
Draft Sequence Produced by DOE Joint Genome Institute  
www.igi.doe.gov  
Finishing Completed at Stanford Human Genome Center  
www.sngc.stanford.edu  
Quality: Phrap Quality >=40 99.9% of Sequence;  
Estimated Total Number of Errors is 0.1.  
FEATURES  
source 1. .141633  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="5"  
/clone="RP11-221L16"

ORIGIN  
Query Match 24.6%; Score 91.2; DB 8; Length 141633;  
Best Local Similarity 71.4%; Pred. No. 5e-18; Indels 7; Gaps 1;  
Matches 137; Conservative 0; Mismatches 48; Indels 7; Gaps 1;

QY 2 GGTTCAGCCACAGCAAAATGTATAGAGTGCTACAAC-----CCTCTATACAGGGATACA 54  
DB 66221 GGTTCAGCCACAGCAAAAGTTTATAAGAGCTACAACAATACTCCAACAATATGAAATCAC 66280  
QY 55 GCAATGTGAACACGAGGTCAATGATCTGAGGTTTCATCAAAACACTGGAAGCCATGGTA 114  
DB 66281 AGACACTTGGTCAGAGGCCAAGCAACTGAGGTTCTGCAAGCTCTGGAAGCCCTGGTA 66340  
QY 115 GTTATTGTCCTTGTATAGGTATGCAACCTCAACCAATGTCCTTCTGGCTCATGGAGA 174  
DB 66341 GTTATTGTCCTTGTATAGGTATGCAACCTCAACCAATGTCCTTCTGGCTCATGGAGA 66400  
QY 175 AATGGGCTGCCC 186  
DB 66401 AAGGATGGCCC 66412

Search completed: December 7, 2005, 04:21:46  
Job time : 3280.23 secs





```
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239933P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241788P.
PR 20-OCT-2000; 2000US-0241789P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250319P.
PR 01-DEC-2000; 2000US-0251160P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
(HUMA-) HUMAN GENOME SCI INC.
XX

XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-541565/60.
XX
XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,
XX useful for preventing, diagnosing and/or treating nervous system cancers
XX and metastases.
XX
XX Disclosure; SEQ ID NO 11655; 1701pp + Sequence Listing; English.
XX
XX The invention relates to novel genes (ABAI1004-ABAI21534) and proteins
XX (ABAI4678-ABAI8001) useful for preventing, treating or ameliorating
XX medical conditions e.g. by protein or gene therapy. The genes are
XX isolated from a range of human tissues disclosed in the specification.
XX The nucleic acids, proteins, antibodies and (ant)agonists are useful in
XX the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
XX ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
XX breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
XX disorders e.g. Addison's disease, allergies, autoimmune haemolytic
XX anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
XX multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
XX cardiovascular disorders such as myocardial ischaemias; (d) wound healing
XX ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
XX infectious diseases such as viral, bacterial, fungal and parasitic
XX infections. Note: The sequence data for this patent did not form part of
XX the printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 27733 BP; 7380 A; 6291 C; 6413 G; 7649 T; 0 U; 0 Other;
XX
XX Query Match 25.2%; Score 93.4; DB 5; Length 27733;
XX Best Local Similarity 80.7%; Pred. No. 2.6e-20;
XX Matches 109; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
XX
XX QY 50 ATACAGCAATGTGAAACGAGAGTCATGTCATGAGGTTTCATCAAAACACTGGAAGCCA 109
XX DB 4609 ATAGTAGACTGTGGCCAGAGGCCATGTCATGAGGTTTGTCCAAACACTGGAAGCCC 4668
XX
XX QY 110 TGGTAGTTATTGTCCTTGATAGTATGTCAACCTCAACACCAATGTCCTCTGCTGAT 169
XX DB 4669 TGGTAGTTATTGTCCTCTTATTAGTATGTCAGCCTCAACACCAATGTCCTCTGCTGAT 4728
XX
XX QY 170 GGAGAAATGGGCTGC 184
XX DB 4729 GGAGAAACAGACAGC 4743
XX
XX RESULT 3
XX ACF87264/c
XX ID ACF87264 standard; DNA; 456 BP.
XX
XX AC ACF87264;
XX
XX DT 02-JUN-2005 (first entry)
XX
XX DE Human SIRS/sepsis diagnostic marker DNA fragment 6124.
XX
XX KW Systemic inflammatory response syndrome; SIRS; antibacterial;
XX immunosuppressive; antiinflammatory; diagnosis; sepsis; ds.
XX
XX OS Homo sapiens.
XX
XX PN WO2004087949-A2.
XX
XX PD 14-OCT-2004.
XX
XX PF 31-MAR-2004; 2004WO-EP003419.
XX
XX PR 02-APR-2003; 2003DE-01015031.
XX PR 08-AUG-2003; 2003DE-01036511.
XX PR 02-SEP-2003; 2003DE-01040395.
XX
```

PA (SIRS-) SIRS LAB GMBH.  
XX Russwurm S, Reinhart K, Saluz H, Straube E, Zipfel PF, Deigner H;  
XX WPI; 2004-748070/73.  
DR In vitro detection of systemic inflammatory response syndrome and related  
XX conditions, for e.g. monitoring progression, comprises detecting abnormal  
XX expression of disease-related genes.  
PT Disclosure; Page; 75pp; German.  
PS  
XX  
XX The invention relates to a novel method for in vitro detection of  
XX systemic inflammatory response syndrome (SIRS). The method comprises  
XX detecting abnormal expression of disease-related genes, or their  
XX associated peptides. The method of the invention demonstrates  
XX antibacterial, immunosuppressive and antiinflammatory applications and  
XX may be used for early differential diagnosis, monitoring progression,  
XX assessing risk, assessing the likely response to treatment and for post  
XX mortem diagnosis of systemic inflammatory response syndrome, sepsis and  
XX sepsis-like conditions. The recombinant or synthetic nucleic acid  
XX sequences of the invention, or derived proteins or peptides, may be  
XX useful as calibrants in assays for the specified diseases, for evaluating  
XX activity or toxicity in screening for active agents and/or for  
XX preparation of agents for treatment or prevention of the specified  
XX diseases. The current sequence is that of a human SIRS/sepsis diagnostic  
XX marker DNA fragment of the invention. Note: The sequence data for this  
XX patent did not form part of the printed specification, but was obtained  
XX in electronic format directly from WIPO at ftp.wipo.int/pub/published  
XX pct sequences. Furthermore, a number of arbitrary SEQ ID NO.s are  
XX disclosed within the specification, however, these have not been taken  
XX into account during indexing due to inconsistencies in application and  
XX format  
XX  
SQ Sequence 456 BP; 107 A; 87 C; 134 G; 123 T; 0 U; 5 Other;  
  
Query Match 24.2%; Score 89.6; DB 13; Length 456;  
Best Local Similarity 67.2%; Pred. No. 7e-20; Indels 0; Gaps 0;  
Matches 125; Conservative 0; Mismatches 61;  
  
QY 31 TACAACCTCTTATACAGGGATACAGCAATGTGAACCCAGAGGTGATGCTGAGGTTTC 90  
DB 236 TGCTCCCAACAATATGAATTTGCCAGANTGTTGCCAGAGGCCACATGCTGAGGTTTC 177  
  
QY 91 ATCAAAACACTGGAGCCATGGTAGTTATTGCTCTGATAGGTATGTCAACCTCAACACC 150  
DB 176 ATGAAGACCTGGAAACCATGGTAGTTACTGTTCTTGAAGTATGTGGCCTCAACACC 117  
  
QY 151 AATGTCCTTCTGCTGATGAGAAATGGGCTGCCCGATTTCACAGACTTCTTGATCTTC 210  
DB 116 AGTGCCCTTCTGCTGATGAGAAATGGATGGCACTTGCCCTGGAAGGATGTCAGCTG 57  
  
QY 211 CTCAGC 216  
DB 56 CCCAGC 51  
  
RESULT 4  
ADH10017\_1/c  
Continuation (2 of 4) of ADH10017 from base 100001 (Human chromosome 2p21-22 fragment cd  
WP Sequence split into 4 fragments LOCUS ADH10017 Accession Adh10017  
WP Fragment Name Begin End  
WP ADH10017\_0 1 110000  
WP ADH10017\_1 100001 210000  
WP ADH10017\_2 200001 310000  
WP ADH10017\_3 300001 365186  
  
Query Match 23.8%; Score 88; DB 10; Length 110000;  
Best Local Similarity 80.5%; Pred. No. 3.7e-18; Indels 0; Gaps 0;  
Matches 103; Conservative 0; Mismatches 25;  
  
QY 243 CTCCTGGCTAGGAGTGGCAGCTTACATCGGTATATCCAGCACCCTTGGGAAGCTGAGGCAG 302

DB 24799 CACCTGGCTGGGCATGGTGGCTCACATCTATAATCCAGCACCTTTGGGAGACCGAGGCAG 24740  
QY 303 GAGGATTACTGGAGCCAGGAGTTTGAGACCAGGCTGGGCAATACAGCGAGACTCTCTCT 362  
DB 24739 GAGGATCACTTGAGCCAGGACTTTTGAGACCAGGCTTTGAGAGACTCTCTCTCTCT 24680  
QY 363 CCTAAAAA 370  
DB 24679 CTACAAAA 24672  
  
RESULT 5  
ADQ97266\_0/c  
WP Sequence split into 4 fragments LOCUS ADQ97266 Accession Adq97266  
WP Fragment Name Begin End  
WP ADQ97266\_0 1 110000  
WP ADQ97266\_1 100001 210000  
WP ADQ97266\_2 200001 310000  
WP ADQ97266\_3 300001 355211  
ID ADQ97266 standard; DNA; 355211 BP.  
XX  
AC ADQ97266;  
XX  
XX 07-OCT-2004 (first entry)  
DT Human cancer associated sequence HD08-023, SEQ ID 242.  
DE  
XX Cytostatic; Gene Therapy; cancer; leukemia; lymphoma; Human; ds.  
KW  
XX Homo sapiens.  
OS  
XX WO2004060304-A2.  
PN  
XX 22-JUL-2004.  
PD  
XX 22-DEC-2003; 2003WO-US041389.  
PF  
XX 27-DEC-2002; 2002US-00330773.  
PR  
XX (SAGR-) SAGRES DISCOVERY INC.  
PA  
XX Morris DW, Malandro MS;  
PI WPI; 2004-543781/52.  
XX  
XX New isolated cancer associated nucleic acids comprising at least 10  
PT contiguous nucleotides, useful for diagnosing, preventing and/or treating  
PT cancers such as leukemia and lymphoma.  
XX  
PS Claim 1; SEQ ID NO 242; 199pp; English.  
XX  
XX The present invention relates to cancer associated sequences (ADQ97025-  
CC ADQ98004). The sequences are useful for the diagnosis, prevention and/or  
CC treatment of cancer, such as leukemia and lymphoma. Note: The sequence  
CC data for this patent did not form part of the printed specification, but  
CC was obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 355211 BP; 96742A; 68188C; 72417G; 106131T; 0U; 117330Other;  
  
Query Match 23.8%; Score 88; DB 12; Length 110000;  
Best Local Similarity 80.5%; Pred. No. 3.7e-18;  
Matches 103; Conservative 0; Mismatches 25; Indels 0; Gaps 0;  
  
QY 243 CTCCTGGCTAGGAGTGGCAGCTTACATCGGTATATCCAGCACCCTTGGGAAGCTGAGGCAG 302  
DB 105102 CACCTGGCTGGGCATGGTGGCTCACATCTATAATCCAGCACCTTTGGGAGACCGAGGCAG 105043  
  
QY 303 GAGGATTACTGGAGCCAGGAGTTTGAGACCAGGCTGGGCAATACAGCGAGACTCTCTCT 362  
DB 105042 GAGGATCACTTGAGCCAGGACTTTTGAGACCAGGCTTTGAGAGACTCTCTCTCTCT 104983  
QY 363 CCTAAAAA 370













PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-02411785P.  
PR 20-OCT-2000; 2000US-02411786P.  
PR 20-OCT-2000; 2000US-02411787P.  
PR 20-OCT-2000; 2000US-02411808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249266P.  
PR 17-NOV-2000; 2000US-0249279P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250150P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0256179P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
XX WPI; 2001-502630/55.  
XX  
XX Polynucleotides encoding digestive system antigens, useful for  
PT diagnosing, treating, preventing and/or prognosing disorders of the  
PT digestive system, particularly cancer and cancer metastases.  
PT  
XX Disclosure; SEQ ID NO 3131; 986pp; English.  
XX  
XX The present invention provides the protein and coding sequences of a  
CC number of human digestive system antigens. These can be used in the  
CC diagnosis, treatment and prevention of digestive system disorders,  
CC including cancer, Meckel's diverticulum, bacterial or parasitic

CC infections, appendicitis, Hirschsprung's disease, chronic colitis or  
CC ulcerative colitis. The present sequence is a genomic DNA fragment  
CC encoding a digestive system antigen of the invention  
XX  
SQ Sequence 23378 BP; 6470 A; 5283 C; 4895 G; 6730 T; 0 U; 0 Other;  
Query Match 23.5%; Score 86.8; DB 4; Length 23378;  
Best Local Similarity 76.8%; Pred. No. 4.4e-18;  
Matches 106; Conservative 0; Mismatches 32; Indels 0; Gaps 0;  
QY 233 TGAACCTCTCTCTGGCTAGGAGTGGCAGCCTACATCCTAATCCAGCACCTTGGGAA 292  
Db 17240 TAAATTTCTCCACTTGGCTGGGTGTGGTCTCACCCCTGTAATCCAGCACCTTGGGAG 17181  
QY 293 GCTGAGGCAGGAGGATTACTGGAGGCCAGGAGTTTGAGACCAGGCTGGGGCAATACAGCGA 352  
Db 17180 GCTGAGGCAGGAGGATCACTAGAGTCCAGGAGTTAGAGACCCAGCTGGGCGCATAGTGA 17121  
QY 353 GACTCTCTCTCTCTAAAAA 370  
Db 17120 GACCTGTCTCTACAAAA 17103  
RESULT 11  
ADB60783/c  
ID ADB60783 standard; DNA; 23378 BP.  
XX ADB60783;  
XX  
DT 04-DEC-2003 (first entry)  
XX  
DE Connective tissue related genomic DNA #526.  
XX  
KW cytostatic; neuroprotective; nootropic; antiparkinsonian; cardiovascular;  
KW antiarteriosclerotic; immunosuppressive; antirheumatic; antiarthritic;  
KW antiinflammatory; antiallergic; antiasthmatic; dermatological;  
KW nephrotropic; virucide; fungicide; antibacterial; antiparasitic;  
KW gene therapy; ds; connective tissues disorder; rheumatoid arthritis;  
KW systemic lupus erythematosus; scleroderma; Sjogren's syndrome; cancer;  
KW cancer metastasis; neoplasmia; leukemia; neurodegenerative disorder;  
KW Alzheimer's disease; Parkinson's disease; cardiovascular disease;  
KW atherosclerosis; myocarditis; cardiopulmonary bypass complication;  
KW autoimmune disease; multiple sclerosis; allergic reaction; asthma;  
KW rhinitis; eczema; inflammatory condition; Crohn's disease; nephritis;  
KW gastrointestinal disorder; inflammatory bowel disease;  
KW organ transplant rejection; immune system disorder; Bruton's disease;  
KW X-linked lymphoproliferative syndrome;  
KW B-cell lymphoproliferative disorder; HIV; AIDS; infection;  
KW chromosome identification; chromosome mapping;  
KW connective tissue related polynucleotide; gene; ds.  
XX Homo sapiens.  
OS  
XX  
XX US2003054375-A1.  
PN  
XX  
XX 20-MAR-2003.  
PD  
XX  
XX 07-MAR-2002; 2002US-00092154.  
PF  
XX  
XX 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184564P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225211P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 18-AUG-2000; 2000US-0225759P.  
PR 22-AUG-2000; 2000US-0226273P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226688P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229511P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231411P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0234844P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.

PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249298P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
PR 17-JAN-2001; 2001US-00764847.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Ruben SM, Barash SC;  
PI  
XX WPI; 2003-634869/60.  
XX P-PSDB; ADB59564.  
XX  
XX New connective tissue-related polypeptides and polynucleotides, useful  
PT for treating, preventing and/or prognosing e.g. disorders of connective  
PT tissue, (e.g. rheumatoid arthritis), cancers, cancer metastases and/or  
PT neoplasias.  
XX  
XX Disclosure; SEQ ID NO 1514; 248pp; English.  
PS  
XX The invention describes an isolated nucleic acid molecule (1), which  
CC comprises a sequence that is at least 95 % identical to a connective  
CC tissue-related polynucleotide encoding connective tissue antigens (CTA).  
CC The polypeptide or polynucleotide is useful for preventing, treating, or  
CC ameliorating medical conditions in a mammal. The connective tissue  
CC polypeptides, polynucleotides and antibodies are particularly useful for  
CC treating, preventing and/or prognosing disorders of connective tissues  
CC (e.g. rheumatoid arthritis, discoid and systemic lupus erythematosus,







CC The present invention relates to novel DNA and protein sequences which  
CC are associated with carcinomas. The sequences are useful for: (i) for  
CC screening drug candidates; (ii) for screening of bioactive agent capable  
CC of binding to carcinoma Associated Protein (CAP); (iii) for screening of  
CC a bioactive agent capable of modulating the activity of CAP; (iv) for  
CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing  
CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating  
CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;  
CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for  
CC determining Carcinoma Associated (CA) gene copy number. In addition, the  
CC CA genes are useful as DNA vaccines and the CAP are useful as markers of  
CC carcinoma including lymphoma. The present sequence is one such CA coding  
CC sequence. Note: This patent is an equivalent to basic patent  
CC US2002182586A1, for which no sequence data was published  
XX  
SQ Sequence 152330 BP; 40094 A; 34646 C; 37146 G; 40111 T; 0 U; 333 Other;  
Query Match 23.1%; Score 85.6; DB 11; Length 152330;  
Best Local Similarity 73.6%; Pred.No. 2.9e-17;  
Matches 109; Conservative 0; Mismatches 39; Indels 0; Gaps 0;  
Qy 223 TACTCTCATATGAACCTCTCTCTGGCTAGGAGTGGCAGCCTACATCGCTAATCCAGC 282  
Db 119498 TAGTACCCCTTATAAGAGAGGCCCGAGGGCTGGCTGGTGGCTCACACCTATAATCCAGC 119439  
Qy 283 ACCTTGGGAAGCTGAGGCAGGAGGATTTACTGGAGGCCAGGAGTTTGAGACCAGGCTGGGC 342  
Db 119438 ACTTTGGGAGGCTGAGGCAGGAGGATTTGTTGGAGCCCGAGGAGTTGGAGACCAGCCTGGGC 119379  
Qy 343 AATACAGCGAGACTCTCTCTCTAAATAA 370  
Db 119378 AGCACAGTGAGACCCCATCTCCACAAA 119351

Search completed: December 7, 2005, 00:02:54  
Job time : 585.669 secs



Db      256 CAACATAGTGAGACCCTGTCTTAAAAA    284

RESULT 2  
US-10-027-632-128851  
; Sequence 128851, Application US/10027632  
; Publication No. US20020198371A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; TITLE OF INVENTION: Polymorphisms in the Human Genome  
; FILE REFERENCE: 108927.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 128851  
; LENGTH: 559  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-128851

Query Match                  24.3%; Score 89.8; DB 5; Length 559;  
Best Local Similarity         75.2%; Pred. No. 3.3e-20;  
Matches 112; Conservative    0; Mismatches 37; Indels 0; Gaps 0;

QY    222 ATACTTCATATGAAACTCTCCTCGTAGGAGTGCCAGCATATCGCTTAATCCCAG 281  
DB    136 AAAATCTCTTTACAAGTCTTTTTCTGGCTGGATGTGGTGCCTCATGACTCTAATTCCAG 195

QY    282 CACCTTGGGAAGCTGAGCAGGAGTACTTGGAGGCCAGAGTTTGAGACCAGGCTGGG 341  
DB    196 CACTTTGGAGGCTGATATGGGAGGATCACTCGAGGCCAGAGTTTGAGACCAGCCTGGG 255

QY    342 CAATACAGCGAGACTCTCTCTCTAAAAA 370  
DB    256 CAACATAGTGAGACCCTGTCTTAAAAA 284

RESULT 4  
US-10-027-632-128850  
; Sequence 128850, Application US/10027632  
; Publication No. US20030204075A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; TITLE OF INVENTION: Polymorphisms in the Human Genome  
; FILE REFERENCE: 108927.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 128850  
; LENGTH: 559  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-128850

Query Match                  24.3%; Score 89.8; DB 6; Length 559;  
Best Local Similarity         75.2%; Pred. No. 3.3e-20;  
Matches 112; Conservative    0; Mismatches 37; Indels 0; Gaps 0;

QY    222 ATACTTCATATGAAACTCTCCTCGTAGGAGTGCCAGCATATCGCTTAATCCCAG 281  
DB    136 AAAATCTCTTTACAAGTCTTTTTCTGGCTGGATGTGGTGCCTCATGACTCTAATTCCAG 195

QY    282 CACCTTGGGAAGCTGAGCAGGAGTACTTGGAGGCCAGAGTTTGAGACCAGGCTGGG 341  
DB    196 CACTTTGGAGGCTGATATGGGAGGATCACTCGAGGCCAGAGTTTGAGACCAGCCTGGG 255

QY    342 CAATACAGCGAGACTCTCTCTCTAAAAA 370  
DB    256 CAACATAGTGAGACCCTGTCTTAAAAA 284

RESULT 3  
US-10-027-632-128852  
; Sequence 128852, Application US/10027632  
; Publication No. US20020198371A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; TITLE OF INVENTION: Polymorphisms in the Human Genome  
; FILE REFERENCE: 108927.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24







FILE REFERENCE: 108827.129  
CURRENT APPLICATION NUMBER: US/10/027,632  
CURRENT FILING DATE: 2002-04-30  
PRIOR APPLICATION NUMBER: US 60/218,006  
PRIOR FILING DATE: 2000-07-12  
PRIOR APPLICATION NUMBER: US 60/198,676  
PRIOR FILING DATE: 2000-04-20  
PRIOR APPLICATION NUMBER: US 60/193,483  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: US 60/185,218  
PRIOR FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/167,363  
PRIOR FILING DATE: 1999-11-23  
PRIOR APPLICATION NUMBER: US 60/156,358  
PRIOR FILING DATE: 1999-09-28  
PRIOR APPLICATION NUMBER: US 60/146,002  
PRIOR FILING DATE: 1999-08-09  
NUMBER OF SEQ ID NOS: 325720  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 325370  
LENGTH: 847  
TYPE: DNA  
ORGANISM: Human  
US-10-027-632-325370

Query Match 23.6%; Score 87.4; DB 5; Length 847;  
Best Local Similarity 74.1%; Pred. No. 2.8e-19;  
Matches 109; Conservative 1; Mismatches 37; Indels 0; Gaps 0;

QY	224	ACTCTCATATGAAACTCTCTCGCTAGGAGTGCGCAGCTACATCGCTAAATCCCGCA	283
Db	140	ACTCTGAAAAGAAATATTAAACAGGCTGGTGGCTCATGCTATATCCAGAA	199
QY	284	CCTTGGGAAGCTGAGGAGGAGATTACTGGAGGCCAGGAGTTTGAGACCAGGCTGGCA	343
Db	200	CTTTGGRGCTGAGGCGAGGATCATCTTGAGGCCAGGAGTTTGAGACCAGGCTGGCA	259
QY	344	ATACAGCGAGACTCTCTCTCTTAAAAA	370
Db	260	ACAAATGAGACTCTGTCTCTAGAAAA	286

RESULT 15  
US-10-027-632-325370  
Sequence 325370, Application US/10027632  
Publication No. US20030204075A9  
GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
TITLE OF INVENTION: Polymorphisms in the Human Genome  
FILE REFERENCE: 108827.129  
CURRENT APPLICATION NUMBER: US/10/027,632  
CURRENT FILING DATE: 2002-04-30  
PRIOR APPLICATION NUMBER: US 60/218,006  
PRIOR FILING DATE: 2000-07-12  
PRIOR APPLICATION NUMBER: US 60/198,676  
PRIOR FILING DATE: 2000-04-20  
PRIOR APPLICATION NUMBER: US 60/193,483  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: US 60/185,218  
PRIOR FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/167,363  
PRIOR FILING DATE: 1999-11-23  
PRIOR APPLICATION NUMBER: US 60/156,358  
PRIOR FILING DATE: 1999-09-28  
PRIOR APPLICATION NUMBER: US 60/146,002  
PRIOR FILING DATE: 1999-08-09  
NUMBER OF SEQ ID NOS: 325720  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 325370  
LENGTH: 847  
TYPE: DNA  
ORGANISM: Human

US-10-027-632-325370

Query Match 23.6%; Score 87.4; DB 6; Length 847;  
Best Local Similarity 74.1%; Pred. No. 2.8e-19;  
Matches 109; Conservative 1; Mismatches 37; Indels 0; Gaps 0;

QY	224	ACTCTCATATGAAACTCTCTCGCTAGGAGTGCGCAGCTACATCGCTAAATCCCGCA	283
Db	140	ACTCTGAAAAGAAATATTAAACAGGCTGGTGGCTCATGCTATATCCAGAA	199
QY	284	CCTTGGGAAGCTGAGGAGGAGATTACTGGAGGCCAGGAGTTTGAGACCAGGCTGGCA	343
Db	200	CTTTGGRGCTGAGGCGAGGATCATCTTGAGGCCAGGAGTTTGAGACCAGGCTGGCA	259
QY	344	ATACAGCGAGACTCTCTCTCTTAAAAA	370
Db	260	ACAAATGAGACTCTGTCTCTAGAAAA	286

Search completed: December 6, 2005, 22:56:44  
Job time : 925.31 secs





QY	306	GATTACTGGAGGCCAGGAGTTTGAGACCCAGGCTGGGGCAATACAGCAGACTCTCTCTCT	365
Db	120315	GATTGCTTTGAGCCAGGAGTTTAAGACCCAGGCTGGGGCAATACAGCAGACTCTCTCTCT	120374
QY	366	AAAAA 370	
Db	120375	CAAAA 120379	
RESULT 4			
US-10-496-711-25			
; Sequence 25, Application US/10496711			
; Publication No. US20050256649A1			
; GENERAL INFORMATION:			
; APPLICANT: SmithKline Beecham Corporation			
; TITLE OF INVENTION: HIGH THROUGHPUT CORRELATION OF			
; TITLE OF INVENTION: POLYMORPHIC FORMS WITH MULTIPLE PHENOTYPES WITHIN CLINICAL			
; TITLE OF INVENTION: POPULATIONS			
; FILE REFERENCE: PU4699WO			
; CURRENT APPLICATION NUMBER: US/10/496,711			
; CURRENT FILING DATE: 2004-05-26			
; PRIOR APPLICATION NUMBER: 60/344892			
; PRIOR FILING DATE: 2002-12-21			
; NUMBER OF SEQ ID NOS: 25			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 25			
; LENGTH: 7240			
; TYPE: DNA			
; ORGANISM: Homo Sapien			
US-10-496-711-25			
Query Match 21.4%; Score 79.2; DB 6; Length 7240;			
Best Local Similarity 77.4%; Pred. No. 5.4e-17;			
Matches 96; Conservative 0; Mismatches 28; Indels 0; Gaps 0;			
QY	247	TGGCTAGGAGTGGCAGCCTACATCGCTAATCCAGCACCTTTGGGAAGCTGAGCAGGAGG	306
Db	2357	TGGCTAGATGAGTGGCTCACACCTGTAATCCAGAACTTTGGAGCTGAGGCAACAGG	2416
QY	307	ATTACTGGAGGCCAGGAGTTTGAGACCCAGGCTGGGCAATACAGCAGACTCTCTCTCT	366
Db	2417	ATTGCTTGAGGACAGGCATTAGACCCAGCCTGAGCAACATAGCAGACTCTCTCTCCAC	2476
QY	367	AAAA 370	
Db	2477	AAAA 2480	
RESULT 5			
US-10-276-233A-6			
; Sequence 6, Application US/10276233A			
; Publication No. US20050260572A1			
; GENERAL INFORMATION:			
; APPLICANT: DNA Chip Research Inc.			
; APPLICANT: Hitachi Software Engineering Co., Ltd.			
; TITLE OF INVENTION: A method of predicting cancer condition			
; FILE REFERENCE: PH-1533-PCT			
; CURRENT APPLICATION NUMBER: US/10/276,233A			
; CURRENT FILING DATE: 2002-11-14			
; PRIOR APPLICATION NUMBER: JP 2001-73063			
; PRIOR FILING DATE: 2001-03-14			
; PRIOR APPLICATION NUMBER: JP 2001-108503			
; PRIOR FILING DATE: 2001-04-06			
; PRIOR APPLICATION NUMBER: JP 2001-234807			
; PRIOR FILING DATE: 2001-08-02			
; NUMBER OF SEQ ID NOS: 27			
; SOFTWARE: PatentIn Ver. 3.2			
; SEQ ID NO 6			
; LENGTH: 79528			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-10-276-233A-6			



Db 95290 CGAGGGGGAGGATTGCTGAGGCCAGGAGTTTGAGACCAGCCTGGGCAACATGGTGAGA 95231

Qy 355 CTCCTCTCTCTAAAA 370  
||||| ||| |||

Db 95230 CTCGTGCTCTACAAA 95215

RESULT 9  
US-10-972-766-1  
; Sequence 1, Application US/10972766  
; Publication No. US20050250118A1  
; GENERAL INFORMATION:  
; APPLICANT: Aerssens, Jeroen  
; APPLICANT: Athanasiou, Maria  
; APPLICANT: Brain, Carlos  
; APPLICANT: Cohen, Nadine  
; APPLICANT: Dain, Bradley  
; APPLICANT: Denton, R. Rex  
; APPLICANT: Judson, Richard  
; APPLICANT: Ozdemir, Vural  
; APPLICANT: Reed, Carol R  
; TITLE OF INVENTION: EPHX2 Genetic Markers Associated with Galantamine Response  
; FILE REFERENCE: 2300.0030001  
; CURRENT APPLICATION NUMBER: US/10/972,766  
; CURRENT FILING DATE: 2004-10-26  
; PRIOR APPLICATION NUMBER: 60/515,378  
; PRIOR FILING DATE: 2003-10-28  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 1  
; LENGTH: 55763  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (11006)..(11006)  
; OTHER INFORMATION: n is the reference allele 'c' which can also be the variant  
; OTHER INFORMATION: allele 't'  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (21845)..(21845)  
; OTHER INFORMATION: n is the reference allele 'a' which can also be the variant  
; OTHER INFORMATION: allele 'g'  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (26376)..(26376)  
; OTHER INFORMATION: n is the reference allele 'g' which can also be the variant  
; OTHER INFORMATION: allele 'a'  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (26434)..(26434)  
; OTHER INFORMATION: n is the reference allele 't' which can also be the variant  
; OTHER INFORMATION: allele 'c'  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (46768)..(46768)  
; OTHER INFORMATION: n is the reference allele 'c' which can also be the variant  
; OTHER INFORMATION: allele 't'  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (54475)..(54475)  
; OTHER INFORMATION: n is the reference allele 'a' which can also be the variant  
; OTHER INFORMATION: allele 'c'  
; US-10-972-766-1  
Query Match 21.1%; Score 78.2; DB 6; Length 55763;  
Best Local Similarity 72.7%; Pred. No. 3.2e-16;  
Matches 101; Conservative 0; Mismatches 38; Indels 0; Gaps 0;  
Qy 232 ATGAACCTCTCTCTCTGCTAGGAGTGGCGACCTTACATCGCTAATCCAGCACCTTTGGGA 291  
Db 24048 ATAAACTTCTCTGTGGGGCCAGGTTGTGGTGCCTCATGCTGTATCCAGCACCTTTGGGA 24107

Qy 292 AGCTGAGGAGGAGGATTACTGGAGCCAGGAGTTTGGAGACAGGCTGGGCAATACAGCG 351  
Db 24108 GGCCAAGGAGGAGGATCACTTAATGCCAGGAGTTTCGAGACCAGCCTGGGCAACATAGTG 24167

Qy 352 AGACTCTCTCTCTCTAAAAA 370  
||||| ||| |||

Db 24168 AGACCCCTCTCTCTAAAAA 24186

RESULT 10  
US-11-112-908-42  
; Sequence 42, Application US/11112908  
; Publication No. US20050260659A1  
; GENERAL INFORMATION:  
; APPLICANT: Harris, Cole  
; APPLICANT: Davis, Lisa M.  
; TITLE OF INVENTION: Breast Cancer Biomarkers  
; FILE REFERENCE: 04-164-US  
; CURRENT APPLICATION NUMBER: US/11/112,908  
; CURRENT FILING DATE: 2005-04-22  
; PRIOR APPLICATION NUMBER: US 60/564,758  
; PRIOR FILING DATE: 2004-04-23  
; PRIOR APPLICATION NUMBER: US 60/575,978  
; PRIOR FILING DATE: 2004-06-01  
; PRIOR APPLICATION NUMBER: US 60/631,702  
; PRIOR FILING DATE: 2004-11-30  
; PRIOR APPLICATION NUMBER: US 60/633,826  
; PRIOR FILING DATE: 2004-12-07  
; NUMBER OF SEQ ID NOS: 511  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 42  
; LENGTH: 155515  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; US-11-112-908-42  
Query Match 21.1%; Score 78.2; DB 7; Length 155515;  
Best Local Similarity 77.2%; Pred. No. 5.3e-16;  
Matches 95; Conservative 0; Mismatches 28; Indels 0; Gaps 0;  
Qy 248 GGCTGAGAGTGGCAGCCTACATCGCTAATCCAGCACCTTTGGGAAAGCTGAGGAGGAGGA 307  
Db 43652 GGCTGGGCTTGGTGGCTCACACCTGTAAATCCAGCACCTTTGGGAAAGCTGAGGAGGAGGA 43711

Qy 308 TTAAGGAGGAGGAGGTTTGGAGCCAGGCTGGGCAATACAGGAGACTCTCTCTCTAA 367  
Db 43712 TCACCTTGAACCTGGGAGTTTGGATCCAGCCTGGGCAACATAGGAGACCTTTGTCTCTAA 43771

Qy 368 AAA 370  
Db 43772 AAA 43774

RESULT 11  
US-11-112-908-43  
; Sequence 43, Application US/11112908  
; Publication No. US20050260659A1  
; GENERAL INFORMATION:  
; APPLICANT: Harris, Cole  
; APPLICANT: Davis, Lisa M.  
; TITLE OF INVENTION: Breast Cancer Biomarkers  
; FILE REFERENCE: 04-164-US  
; CURRENT APPLICATION NUMBER: US/11/112,908  
; CURRENT FILING DATE: 2005-04-22  
; PRIOR APPLICATION NUMBER: US 60/564,758  
; PRIOR FILING DATE: 2004-04-23  
; PRIOR APPLICATION NUMBER: US 60/575,978  
; PRIOR FILING DATE: 2004-06-01  
; PRIOR APPLICATION NUMBER: US 60/631,702  
; PRIOR FILING DATE: 2004-11-30  
; PRIOR APPLICATION NUMBER: US 60/633,826  
; PRIOR FILING DATE: 2004-12-07  
; NUMBER OF SEQ ID NOS: 511





GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 21:43:03 ; Search time 2600.1 Seconds  
(without alignments)  
6427.437 Million cell updates/sec

Title: US-09-980-046B-2  
Perfect score: 294  
Sequence: 1 cggccttaaggttcctctga.....aagagagtaattcccaaaa 294

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues  
Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

1: gb\_ba.\*  
2: gb\_in.\*  
3: gb\_env.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pr.\*  
9: gb\_ro.\*  
10: gb\_sts.\*  
11: gb\_sy.\*  
12: gb\_un.\*  
13: gb\_vi.\*  
14: gb\_hg.\*  
15: gb\_pl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	286.6	97.5	163338	8	AL162426 Human DNA
2	286.6	97.5	172588	14	AC007430 Homo sapi
3	49.4	16.8	182870	2	AC116960 Dictyoste
4	49	16.7	110000	14	Continuation (13 o
5	49	16.7	110000	14	Continuation (14 o
6	48.8	16.6	105682	2	Continuation (4 of
7	48.6	16.5	155573	14	CR381703 Danio rer
8	48.4	16.5	188883	14	CR933758 Danio rer
9	47.8	16.3	80518	14	AL109815 Plasmodiu
10	47.2	16.1	15714	6	AX346075 Sequence
11	47.2	16.1	15714	6	AX458542 Sequence
12	47	16.0	152913	5	BX322651 Zebrafish
13	47	16.0	186497	14	CR933524 Danio rer
14	46.8	15.9	138046	14	AC117987 Felis cat
15	46.8	15.9	165105	5	AL929386 Zebrafish
16	46.4	15.8	110000	2	Continuation (5 of
17	46.4	15.8	160167	9	AC107810 Mus muscu
18	46.4	15.8	242741	14	AC106279 Rattus no

c 19	46.2	15.7	256879	2	AC116982
c 20	46	15.6	52838	14	AC087538
c 21	46	15.6	93307	8	AL359737
c 22	45.8	15.6	24152	15	AY962591
c 23	45.8	15.6	160491	14	CR387984
c 24	45.8	15.6	161532	8	AC146202
c 25	45.8	15.6	188132	9	AC129290
c 26	45.6	15.5	1321	15	AF120715
c 27	45.6	15.5	250421	2	AE014849
c 28	45.4	15.4	123205	9	AL933155
c 29	45.4	15.4	144449	14	AC067926
c 30	45.4	15.4	162632	8	AC093916
c 31	45.2	15.4	2664	2	AF206306
c 32	45.2	15.4	112714	5	BX511189
c 33	45.2	15.4	153212	14	CR936952
c 34	45.2	15.4	270106	14	CR392327
c 35	44.8	15.2	14924	6	AX281279
c 36	44.8	15.2	14924	6	AX345126
c 37	44.8	15.2	110000	14	PFMAL8P1_04
c 38	44.8	15.2	176186	8	AC006925
c 39	44.8	15.2	250029	2	AB014816
c 40	44.8	15.2	333321	2	AC116986
c 41	44.6	15.2	1121	2	LMU00101
c 42	44.6	15.2	1141	2	KFLMDIVRA
c 43	44.6	15.2	123687	4	CR450381
c 44	44.6	15.2	154065	5	CR854838
c 45	44.4	15.1	3488	2	AP447563

#### ALIGNMENTS

AL162426 163338 bp DNA linear PRI 18-MAY-2005  
Human DNA sequence from clone RP11-56D16 on chromosome 9 Contains  
the STXBPI gene for syntaxin binding protein 1  
(UNC18, HUNC18, MUNC18-1), two novel genes, a novel gene (LOC138428),  
the gene for a novel protein (LOC286207), the gene for anovel  
protein (FLJ32780), the TOR2A gene for torsin family 2, member A  
(FLJ14771), the 3' end of the SH2D3C gene for SH2 domain containing  
3C (CHAT.NSP3) and five CpG islands, complete sequence.

AL162426 GI:13624960  
HTG; CHAT; CPG Island; FLJ14771; FLJ32780; HUNC18; LOC138428;  
LOC286207; MUNC18-1; NSP3; SH2D3C; STXBPI; TOR2A; UNC18.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.

1 (bases 1 to 163338)  
Kimberley, A.  
Direct Submission  
Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk

Clone requests: clonerequest@sanger.ac.uk  
On Apr 14, 2001 this sequence version replaced gi:13396434.

The following abbreviations are used to associate primary accession  
numbers given in the feature table with their source databases:  
Em, EMBL; Sw, SWISSPROT; Tr, TrEMBL; Wp, WormPEP; Information  
on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C\_elegans/wormpep This sequence  
was generated from part of bacterial clone contigs of human  
chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping  
Group. Further information can be found at

http://www.sanger.ac.uk/HGP/Chr9  
RP11-56D16 is from the library RPI-11.1 constructed by the group  
of Pieter de Jong. For further details see  
http://www.chori.org/bacpac/home.htm

VECTOR: pBACe3.6  
----- Genome Center  
Center: Wellcome Trust Sanger Institute









```
2.71, 5.62 - GSCJ_ID dd_00996"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAO53118.1"
/db_xref="GI:28850333"
/translation="MELEKVEKDLLELVGVDPDIPSELISFGNESSWLNTNSDKF
KTFLN"
complement(join(9304..9432,9619..10184,10260..12186,
12335..12463))
/note="GeneID exon scores (in order of location ranges):
2.33, 62.67, 203.85, 8.24 - GSCJ_ID dd_00995"
/codon_start=1
/product="similar to Homo sapiens (Human). DNA mismatch
repair protein Msh2"
/protein_id="AAO53119.1"
/db_xref="GI:28850334"
/translation="MSDNEORESSOVVLKEDKTFVTFQSLVSNEDDTLRLPDKK
SLKWSIDPDKPKKIKIDNGLSTTTASSQQOQELGLAVLTIQGYEYENIVKEL
LDEKKKIISWKNPSKOQWELIKKSGNTQMFEDVLNGNCEGVMALKVTKRG
SIVGIFGDATFTKIGVSQFMDNLSLSSFMQSVKELCCDQKNDYQYKVE
KLSNADIGPTELKPSFKNAEQDLTLRLASVKNLPDIQEHAIQAGSCLIKHLDL
LSPNPFKFKLEKLDLRYMKLIDSSSFKGLHIIDLKSSVSAAGGAGGAGSSSN
KDQSLYNLNOCTPMGSRLLLOVVKQPLNAREELEARNFEAFYNDELRQSLRSN
DKLKGIDRLSKLHGKATLEDCLNLYGIVTRLPVVLQSLNHHSSHOBLIKVFI
ESLESIISDFKCAWKEITDLDLNDKHEYVIRSEFDETLRGILKKQDOISNKIER
FRVDTADLNDLDEAKVKLHYSEKDFLLRISKDEVAIRDKKYIVHATAKDGVRFAT
REIDTLNAYKWSAEYDKQDLAKRLQIAAASFVPLIEDLSLIATLDVFTLSHV
SSIAPIPTFRPIIPLGSDENGATVIIGGRHPCVEIQDNVNFANDLDRGOSQFQ
ITGPMGSGFTPIRQVGLVMAOIGCFVPAQKATIAVDCILSRVAGSOLRGVS
TFMAELTSYILKVATKNSLIIDELRGSTSYDGLGAWGIABYICNQIGGCLFA
THFHELTLSLDLPMVKNLHVSASTNNMTFTLLYKVEQPCQSPFGIHAVILLANFSPQ
VIENAKQAKALESEFESNTLKQNHKFLIEEFKFNFNFSNDVKSLSVLNLSLYSID
IN"
complement(join(12839..13000,13400..13435))
/note="GeneID exon scores (in order of location ranges):
2.76, -2.24 - GSCJ_ID dd_03346"
/codon_start=1
/product="similar to Plasmodium falciparum. Hypothetical
protein"
/protein_id="AAO53120.1"
/db_xref="GI:28850335"
/translation="MLSNGLKIDQEKPFILLIIINNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNN"
14458..15615
/note="GeneID exon scores (in order of location ranges):
104.42 - GSCJ_ID dd_00993"
/codon_start=1
/product="similar to Mus musculus (Mouse). Similar to
S-adenosylmethionine decarboxylase 1"
/protein_id="AAO8497.2"
/db_xref="GI:28850336"
/translation="WENNNDYKRSTEIFNEADAGFVYQSNDSYDGFEGPEKKLDI
RFGPTSKGSVKSIAGSPKVLRTIDKEKQTVLDSARCTIISQTSNDHMSYVLSES
SLFVYRRAMIKTCGTITLHLIAKQVKGCGLEVEWFSRKNLQPKSQVFPHC
SFSDEYFLNKLFDQAYVMGVNDKWNLYIADFRKNPQLQRTQTEFEVMMHLDDET
VMKQFIRGVSAMDTTVNSGADLLPGSMIDDFQPCGYSMGLLAEFYFTIITP
ESHCSVSDPTVALVALLAKVLNFKPGRFTAALVAEDGAPCGDPYAFDVNVP
SVAIQNTVHGPDGVDVVSNTYQLDKKVNFSNDLLQSIESTNIDLIQV"
complement(16030..17082)
/note="GeneID exon scores (in order of location ranges):
70.61 - GSCJ_ID dd_00992"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAO8498.2"
/db_xref="GI:28850337"
/translation="MKCFILNIIKRNIIKNNIIFKASYPHSPFNNSNIIRNFSSNNKI
NNSNNKEDLMIIPSRSLSEKLSGLYLPNEEYEEVDFDEEYEDENDDDNNN
NNNNNIIKEEKEEKEVENQNLAKGFIKGGKKFQDMYIRIKENHLSNKHSDIFND
VGEQMSFEIERKVEERLAKYQEKWEILDSMDQSDFIYKDNEKVDLYNKIISNFKA
DTHFTFPIITKQLENQLIKPNSGLSMTSNDSDSAFNKYQSDDISRSIKRIQRIQ
KOKLTNHRINGENLVDFSLYREDPILNNAQFLAKFLNLSKSESVENLLKYITPIIVK
FNTGIKSGHNVIIET"
join(18407..18484,18580..21294)
```

```
/note="GeneID exon scores (in order of location ranges):
7.28, 279.86 - GSCJ_ID dd_03068"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAO53121.1"
/db_xref="GI:28850338"
/translation="MDKYDTILWELRELENSKCADCTDSFPRYMTTYGTVCVSCG
AIHRELGNRVKSISSDXTQODIERLEKVGKNMADEIWLKSWQQOYELPFPFSDKRV
RDFLTKYIEKKWIKEGIKOSDFISPTQSTPGTPTTSPKQQQQQPSAVSQKQT
TSPLRNGEKKELITSKDSLSEKFEISLGNLNLNTNQOQQQQQQQQQQQQQSTQ
TSQOISPIQABINQQOQQQQQFQDNGKFPNPRDENKYFNSKPPSIYDIOYKQOQ
SEKPNHSLYNEKSTWADQSLGNSSFYKIOVHPTRKSPSTSIYDIOYKQOQ
QQOQQQQQKQYQONQOQKQLQPOQSOYQQOQQQPPYVINOQPOYNQNTSMQ
NQMGHANQMNQMNQMNQMNQMNQMNQMNQMNQMNQMNQMNQMNQMNQMNQMN
NQLALVPVEQNNNNNNNNNNQVYPPQPPQOHNMFQOQHQPOQOQQOQQLSN
IMNNNSNIENGPFGLISQVLHQHQQMNQMNQMNQMNQMNQMNQMNQMNQMN
PATTTTTIHTPGLNPLSSSSSSSGNFDQOQOQOQPLDMEFKVNSDPVFPAS
SSSPESGNLSDRSEKQEWRRIOEKNIQIQNDKEIARKQLBEEIPIKTRPRNTE
PSPRDRTSLQLOQSQPOQOQYHYSSSYSSDYDTANNHGNPITQOHIVSGHRS
SSGNLTLRLSGRSRSKSYDDPFEPINPRPLNKSQGGYDDSDFMKKNKRIPGGGSNY
QPKISQDQYQDNNNNNNNNNNYQQQSPSOQIVVHQSNYNNNNNNNNNNNANRYSNECD
YCGKDVGLSEMSYHKNTCEILRQBRCSMCKLVKHMQMDHQLQOCTNTYTCGCGRL"
Query Match 16.8%; Score 49.4; DB 2; Length 182870;
Best Local Similarity 49.8%; Pred. No. 3.7; Mismatches 0; Gaps 0;
Matches 125; Conservative 0; Indels 126;

Qy 44 TTTGAGCATCTTACTTCAAATATTCTTCCAGTGTTTGTGTTTCTCAATGTTAT 103
Db 58468 TTTGATCTTACACAATAATAATCATGATTTTAAAAGAAGTTCATATTTTCTAAATTTA 58527

Qy 104 AGTGAATATATATATATACATGAGAAAAGCATAAATCAACTATTGAATACGAAATG 163
Db 58528 GACTAAAAATCATTTTACAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 58587

Qy 164 TGAACCTCTATGTAACGCACTCAGGCTGCTACCGTTTTTAAATTTCTCTCAATCTTT 223
Db 58588 ATAATAGAGATCTCAATTTAGTAATCTCAGCCCATTTTTTTTTTTTTTTTTTAAATCCCA 58647

Qy 224 AAAAAAATACTATCTTTTACTAAAGATGTTTTCAGAGTAAGATAAAATACAAAGAGTA 283
Db 58648 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATA 58707

Qy 284 ATTCCCAAAA 294
Db 58708 AAAATAAAAAA 58718

RESULT 4
PFMAL13_12
WPCOMMENT
Sequence split into 28 fragments LOCUS PFMAL13 Accession AL844509
Fragment Name Begin End
PFMAL13_00 1 110000
PFMAL13_01 100001 210000
PFMAL13_02 200001 310000
PFMAL13_03 300001 410000
PFMAL13_04 400001 510000
PFMAL13_05 500001 610000
PFMAL13_06 600001 710000
PFMAL13_07 700001 810000
PFMAL13_08 800001 910000
PFMAL13_09 900001 1010000
PFMAL13_10 1000001 1110000
PFMAL13_11 1100001 1210000
PFMAL13_12 1200001 1310000
PFMAL13_13 1300001 1410000
PFMAL13_14 1400001 1510000
PFMAL13_15 1500001 1610000
PFMAL13_16 1600001 1710000
PFMAL13_17 1700001 1810000
PFMAL13_18 1800001 1910000
PFMAL13_19 1900001 2010000
```



```

COMMENT      On Mar 19, 2005 this sequence version replaced gi:46879102.
-----
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zfish-help@sanger.ac.uk
-----
Center project name: zK72P23
-----
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 154326 bases at least Q40
Consensus quality: 154612 bases at least Q30
Consensus quality: 154899 bases at least Q20
Insert size: 155273; sum-of-contigs
Insert size: 156492; 2.2% error; agarose-fp
Quality coverage: 10.08x in Q20 bases; sum-of-contigs Quality
coverage: 10.00x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 10164: contig of 10164 bp in length
* 10165 10264: gap of 100 bp
* 10265 33043: contig of 22779 bp in length
* 33044 33143: gap of 100 bp
* 33144 51452: contig of 18309 bp in length
* 51453 51552: gap of 100 bp
* 51553 155573: contig of 104021 bp in length.
-----
FEATURES             Location/Qualifiers
source               1..155573
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /clone="DKEX-72P23"
                     /clone_lib="Daniokey"
misc_feature         1..10164
                     /note="assembly_fragment:00010
                     fragment_chain:1"
misc_feature         10265..33043
                     /note="assembly_fragment:00564
                     fragment_chain:1"
misc_feature         33144..51452
                     /note="assembly_fragment:00195.0"
misc_feature         51553..155573
                     /note="assembly_fragment:01026"
ORIGIN
Query Match      16.5%; Score 48.6; DB 14; Length 155573;
Best Local Similarity 51.2%; Pred. No. 5.4;
Matches 105; Conservative 0; Mismatches 100; Indels 0; Gaps 0;
QY 51 ATCTCTTACTTCAAATATTCCTCCAGTGGTTTGTATCTTCTCAATGTTATAGTAAA 110
Db 155368 ATATATATATTTAAAAATTTTTTTTTTTTTTTTTTTTTTTTANNTTTTTNANNAAT 155427
QY 111 TATAATAATACATGCAGAAAAAGCAATAATCAACTATTGATAACGAAAATGTGAACCT 170
Db 155428 AAAAAAATAAAAAAATAAAAAAATAATAATTAATAATAAAAAAATAAAAAA 155487
QY 171 CTATGTAACTGCAACTCAGCGTCGCTTACCGTTTAAATTTCTCTTAATCTTTAAAAACA 230
Db 155488 AATAAAAAATAAAAAAATTTATTTTATATTTTTTTTTTATAATAAAAAAATAAAAA 155547
QY 231 AATACTATCTTTACTAAAGATGTTT 255
Db 155548 AAAAAATATTTTTTTTAAATATTTTT 155572

```

---

```

RESULT 8
CR933758 188883 bp DNA linear HTG 01-JUN-2005
LOCUS    Danio rerio clone CH211-203E21, WORKING DRAFT SEQUENCE, 3 unordered
DEFINITION
CR933758.3 GI:66863493
VERSION  HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.
KEYWORDS  Danio rerio (zebrafish)
SOURCE    Danio rerio
ORGANISM  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
          Cypriniformes; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 188883)
AUTHORS   McLaren,S.
TITLE     Direct Submission
JOURNAL   Submitted (30-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
          Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
          zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
          On Jun 1, 2005 this sequence version replaced gi:58197316.
COMMENT   -----
          Genome Center
          Center: Wellcome Trust Sanger Institute
          Center code: SC
          Web site: http://www.sanger.ac.uk
          Contact: zfish-help@sanger.ac.uk
          -----
          Project Information
          Center project name: zC203E21
          -----
          Summary Statistics
          Assembly program: XGAP4; version 4.5
          Chemistry: Dye-terminator; 100% of reads
          Consensus quality: 187341 bases at least Q40
          Consensus quality: 187488 bases at least Q30
          Consensus quality: 187715 bases at least Q20
          Insert size: 188683; sum-of-contigs
          Insert size: 194743; 3.4% error; agarose-fp
          Quality coverage: 9.20x in Q20 bases; sum-of-contigs Quality
          coverage: 9.05x in Q20 bases; agarose-fp
          -----
          * NOTE: This is a 'working draft' sequence. It currently
          * consists of 3 contigs. The true order of the pieces
          * is not known and their order in this sequence record is
          * arbitrary. Gaps between the contigs are represented as
          * runs of N, but the exact sizes of the gaps are unknown.
          * This record will be updated with the finished sequence
          * as soon as it is available and the accession number will
          * be preserved.
          *
          * 1 90306: contig of 90306 bp in length
          * 90307 90406: gap of 100 bp
          * 90407 98661: contig of 8255 bp in length
          * 98662 98761: gap of 100 bp
          * 98762 188883: contig of 90122 bp in length.
          -----
          Location/Qualifiers
          source       1..188883
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /clone="CH211-203E21"
                     /clone_lib="CHORI-211"
                     /clone_size="left"
                     /note="assembly_fragment:01266
                     fragment_chain:1
                     clone end:SP6
                     vector_side:left"
                     90407..98661
                     /note="assembly_fragment:00022
                     fragment_chain:1"
                     98762..188883
                     /note="assembly_fragment:00108
                     fragment_chain:1"
          misc_feature 1..90306
                     /note="assembly_fragment:01266
                     fragment_chain:1
                     clone end:SP6
                     vector_side:left"
                     90407..98661
                     /note="assembly_fragment:00022
                     fragment_chain:1"
                     98762..188883
                     /note="assembly_fragment:00108
                     fragment_chain:1"
          misc_feature 1..90306
                     /note="assembly_fragment:01266
                     fragment_chain:1
                     clone end:SP6
                     vector_side:left"
                     90407..98661
                     /note="assembly_fragment:00022
                     fragment_chain:1"
                     98762..188883
                     /note="assembly_fragment:00108
                     fragment_chain:1"
          misc_feature 1..90306
                     /note="assembly_fragment:01266
                     fragment_chain:1
                     clone end:SP6
                     vector_side:left"
                     90407..98661
                     /note="assembly_fragment:00022
                     fragment_chain:1"
                     98762..188883
                     /note="assembly_fragment:00108
                     fragment_chain:1"
          ORIGIN
Query Match      16.5%; Score 48.4; DB 14; Length 188883;

```



```
Db 6485 CCTTCCAACTTAAATATATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 6426
Qy 151 AATAAGGAATGTAACCTCTATGTAAGTCACTCAGGCTGCTCGCTTTTAAATTT 210
Db 6425 AAAATCCCAAAAAAATATATATAAAAAAAAAAACTCAAACTAACTA 6373
Qy 211 CTCTCTAATCTTTAAAAACAATACTATCTTTTACTTAAAGATGTTTCAGAGTAAGATAAAA 270
Db 6372 TTCCCTCATATTAATAACAATACCATATTAACCTAAATTTAATTCACAATAACATCAA 6313
Qy 271 TACA 274
Db 6312 AAAA 6309

RESULT 11
AX458542/c
LOCUS AX458542 15714 bp DNA linear PAT 08-JUL-2002
DEFINITION Sequence 88 from Patent W00246454.
ACCESSION AX458542
VERSION AX458542.1 GI:21725206
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Schacht, O.
TITLE Diagnosis of diseases associated with angiogenesis
JOURNAL Patent: WO 0246454-A 88 13-JUN-2002;
Epigenomics AG (DE)
FEATURES
source
1. .15714
Location/Qualifiers
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="chemically treated genomic DNA (Homo sapiens)"

ORIGIN
Query Match 16.1%; Score 47.2; DB 6; Length 15714;
Best Local Similarity 52.9%; Pred. No. 22;
Matches 129; Conservative 0; Mismatches 108; Indels 7; Gaps 1;

Qy 31 AAGAAGATCAGGATTTGAGCATCTCTTACTTCAAAATTTCTTCCAGTGGTTTGTATC 90
Db 6545 ACGAAATCCAACTTCGCTACTTATTTCTCTTTATTTCTTCTCATCCCAATC 6486
Qy 91 TTCTCAATGTTTATAGTGAATATAATAATACATGCAGAAAAAAGCATAAATCAACTATG 150
Db 6485 CCTTCCAACTTAAATTTATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 6426
Qy 151 AATAAGGAATGTAACCTCTATGTAAGTCACTCAGGCTGCTCGCTTTTAAATTT 210
Db 6425 AAAATCCCAAAAAAATATATATAAAAAAAAAAACTCAAACTAACTA 6373
Qy 211 CTCTCTAATCTTTAAAAACAATACTATCTTTTACTTAAAGATGTTTCAGAGTAAGATAAAA 270
Db 6372 TTCCCTCATATTAATAACAATACCATATTAACCTAAATTTAATTCACAATAACATCAA 6313
Qy 271 TACA 274
Db 6312 AAAA 6309

RESULT 12
BX322651
LOCUS BX322651 152913 bp DNA linear VRT 29-JAN-2004
DEFINITION Zebrafish DNA sequence from clone DKEYP-32G10, complete sequence.
ACCESSION BX322651
VERSION BX322651.3 GI:41392279
KEYWORDS HTG.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
```

```
REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT
```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.

1 (bases 1 to 152913)

Direct Submission

Submitted (14-JAN-2004) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk

On Jan 29, 2004 this sequence version replaced gi:30387034.

----- Genome Center

Center: Wellcome Trust Sanger Institute

Center code: SC

Web site: <http://www.sanger.ac.uk>

Contact: [zfish-help@sanger.ac.uk](mailto:zfish-help@sanger.ac.uk)

-----

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at [http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) Clone-derived Zebrafish pUC subclones occasionally display inconsistency over the length of mononucleotide A/T runs and conserved TA repeats. Where this is found the longest good quality representation will be submitted.

Repeat names beginning 'Dr' were identified by the Recon repeat discovery system (Zhirong Bao and Sean Eddy, submitted), and those beginning 'drr' were identified by Rick Waterman (Stephen Johnson lab, WashU). For further information see [http://www.sanger.ac.uk/Projects/D\\_rerio/fishmask.shtml](http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml) DKEYP-32G10 is from a Zebrafish BAC library

VECTOR: pIndigoBAC-5

#### FEATURES

source

Location/Qualifiers

1. .152913

/organism="Danio rerio"

/mol\_type="genomic DNA"

/db\_xref="taxon:7955"

/clone="DKEYP-32G10"

/clone\_lib="DanioKeyPilot"

#### ORIGIN

Query Match 16.0%; Score 47; DB 5; Length 152913;

Best Local Similarity 59.3%; Pred. No. 11;

Matches 80; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

Qy 43 ATTTGAGCATCTCTTACTTCAAAATATTCTTCCAGTGGTTTTTGTATCTTCTCAATGTA 102

Db 56862 ATTGACCACAGTGTGAAGCATCTCTTTTTCCTGTATCATCTTTCTCTTATGTCA 56921

Qy 103 TAGTGAATATATAATACATGCAGAAAAAGCAATAATCAACTATTGAATACCAAAAT 162

Db 56922 CAGTTAATTGAAAAAACTTAGCAAAACAACTAATTAATATCANTAGCAAAAT 56981

Qy 163 GTGAACCTCTATGTA 177

Db 56982 GTAAACCTTTATGTA 56996

RESULT 13









## COMMENT

Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
 zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk  
 On Sep 23, 2003 this sequence version replaced gi:33386428.  
 ----- Genome Center  
 Center: Wellcome Trust Sanger Institute  
 Center code: SC  
 Web site: <http://www.sanger.ac.uk>  
 Contact: zfish-help@sanger.ac.uk  
 -----

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at [http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) Clone-derived Zebrafish pUC subclones occasionally display inconsistency over the length of mononucleotide A/T runs and conserved TA repeats. Where this is found the longest good quality representation will be submitted.

Repeat names beginning 'Dr' were identified by the Recon repeat discovery system (Zhirong Bao and Sean Eddy, submitted), and those beginning 'dr' were identified by Rick Waterman (Stephen Johnson lab, WashU). For further information see [http://www.sanger.ac.uk/Projects/D\\_reio/fishmask.shtml](http://www.sanger.ac.uk/Projects/D_reio/fishmask.shtml) CH211-69J13 is from a CHORI-211 BAC library  
 VECTOR: pTARBAC2.1.

## FEATURES

## source

1. 165105  
 /organism="Danio rerio"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:7955"  
 /clone="CH211-69J13"  
 /clone\_lib="CHORI-211"

## ORIGIN

Query Match	15.9%	Score 46.8;	DB 5;	Length 165105;
Best Local Similarity	50.9%	Fred. No. 11;		
Matches 111;	Conservative 0;	Mismatches 107;	Indels 0;	Gaps 0;

  

Qy	62	CAAAATTATCTTCACAGTGGTTTGGTATCTCTCAATGTTATAGTGAAATATAATAATAC	121
Db	155302	CAATTAAATAAGACTACAGTGTTTATATATACATGTTTAAATGATAAACCAATATAA	155243
Qy	122	ATGCAGAAAAAGCATAAATCAACTATTGGAATAACGAAATGTGAACCTCTATGTAACGTG	181
Db	155242	ATTTTGTACTAGGAAATATGAACCTTTTGAATAAAGGACATAAATTCACATGCAATAAC	155183
Qy	182	CAACTCAGCGTCCTACCGTTTAAATTTCTCTCTAATCTTTAAACAAATACTATCTT	241
Db	155182	ATATTCAATTAAATAGCTTTTCATACACTATCTCTACAAACGTTGACAAAAAGCACATGTT	155123
Qy	242	TACTAAAGATGTTTCAGAGTAGAATAAATAACAAGAG	279
Db	155122	TTTTTGTGTTAGATTGAAGATTAGGGCTACAG	155085

Search completed: December 7, 2005, 04:21:51  
 Job time : 2605.1 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 11:47:47 ; Search time 462.191 Seconds  
(without alignments)  
4239.411 Million cell updates/sec

Title: US-09-980-046B-2  
Perfect score: 294  
Sequence: 1 cggccttaaggttcctctga.....aagagagtaattcccaaaaa 294

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq 21.\*

1: Geneseqn1980s.\*  
2: Geneseqn1990s.\*  
3: Geneseqn2000s.\*  
4: Geneseqn2001as.\*  
5: Geneseqn2001bs.\*  
6: Geneseqn2002as.\*  
7: Geneseqn2002bs.\*  
8: Geneseqn2003as.\*  
9: Geneseqn2003bs.\*  
10: Geneseqn2003cs.\*  
11: Geneseqn2003ds.\*  
12: Geneseqn2004as.\*  
13: Geneseqn2004bs.\*  
14: Geneseqn2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	294	100.0	294	4 AAC89733	AAC89733 Human gas
2	288.2	98.0	100445	13 ABD33179	ABD33179 Human can
3	47.2	16.1	15714	6 ABL33173	ABL33173 Human imm
4	47.2	16.1	15714	6 ABQ67058	ABQ67058 Human ang
5	44.8	15.2	14924	6 ABL32224	ABL32224 Human imm
6	44.8	15.2	14924	6 ABL54321	ABL54321 Chemical
7	44	15.0	5311	6 ABL33019	ABL33019 Human imm
8	43.8	14.9	6197	6 ABN80257	ABN80257 Human che
9	43.8	14.9	34548	6 ABL70604	ABL70604 Chemical
10	43.6	14.8	510	12 ACH78930	Ach78930 Human gen
11	43.4	14.8	5204	6 ABL32901	ABL32901 Human imm
12	43.4	14.8	8170	6 ABK28258	ABK28258 DNA trans
13	43.4	14.8	8423	6 ABL33407	ABL33407 Human imm
14	43.2	14.7	14551	6 ABL34584	ABL34584 Human met
15	43.2	14.7	14551	7 ADS99845	Ades99845 Bisulphit
16	43	14.6	9347	6 ABL49336	ABL49336 Human pol
17	43	14.6	110000	2 AAU21209_04	Continuation (5 of
18	42.6	14.5	5980	6 ABL33189	ABL33189 Human imm
19	42.6	14.5	6274	4 AAS46322	Aas46322 Tumour su

c 20	42.6	14.5	6274	6 ABK33952	Abk33952 Human DNA
c 21	42.6	14.5	6274	6 ABK31239	ABK31239 Signal tr
c 22	42.6	14.5	6274	8 ADA20355	Ada20355 Prostate
c 23	42.6	14.5	6274	8 ADA84162	Ada84162 Human ren
c 24	42.6	14.5	6274	10 ADB54250	Adb54250 Pretreate
c 25	42.6	14.5	6274	10 ADB54122	Adb54122 Pretreate
c 26	42.2	14.4	5163	6 ABL33249	ABL33249 Human imm
c 27	42.2	14.4	7008	4 AAS46531	Aas46531 Tumour su
c 28	42.2	14.4	9289	10 ADB84033	Adb84033 5' regula
c 29	41.8	14.2	8873	6 ABK31211	Abk31211 Signal tr
c 30	41.8	14.2	8873	6 ABL70174	ABL70174 Chemical
c 31	41.8	14.2	8873	6 AAS61125	Aas61125 Human gen
c 32	41.8	14.2	11178	6 ABK31172	ABK31172 Signal tr
c 33	41.8	14.2	11178	6 ABL70507	ABL70507 Chemical
c 34	41.8	14.2	11178	6 AAS61059	Aas61059 Human gen
c 35	41.6	14.1	6104	6 ABL33125	ABL33125 Human imm
c 36	41.6	14.1	7829	6 ABL33105	ABL33105 Human imm
c 37	41.6	14.1	7829	6 ABK31287	ABK31287 Signal tr
c 38	41.6	14.1	7829	6 ABL70266	ABL70266 Chemical
c 39	41.6	14.1	7829	6 AAS61191	Aas61191 Human gen
c 40	41.6	14.1	8056	8 ABZ10246	Abz10246 Haematopo
c 41	41.6	14.1	16579	10 ADB54245	Adb54245 Pretreate
c 42	41.6	14.1	16579	10 ADE37772	Ades37772 Human che
c 43	41.6	14.1	16579	13 ADS89547	Ades89547 Oligonuc
c 44	41.4	14.1	3500	8 ACF62826	ACF62826 Colon can
c 45	41.4	14.1	6222	6 ABL32692	ABL32692 Human imm

## ALIGNMENTS

## RESULT 1

AAC89733

ID AAC89733 standard; cDNA; 294 BP.

AC AAC89733;

DT 12-MAR-2001 (first entry)

DE Human gastrointestinal inflammation-related cDNA, SEQ ID NO: 2.

KW Human; cytostatic; immunomodulator; immunostimulant; vulnery;

KW anti-inflammatory; neuroprotective; antibacterial; gene therapy;

KW gastrointestinal inflammation; immune system disorder; genetic disorder;

KW cancer; autoimmune disorder; infection; wound healing; ss.

OS Homo sapiens.

XX WO200073324-A2.

PN 07-DEC-2000.

PF 01-JUN-2000; 2000WO-US015191.

PR 01-JUN-1999; 99US-0137058P.

PA (DIGI-) DIGITAL GENE TECHNOLOGIES INC.

PI Youakim A, Dubose RF, Sims JE, Pribyl TM, Hillbush BS, Hasel KW;

XX WPI; 2001-061508/07.

XX New polynucleotides and polypeptides, useful in gene therapy and in diagnosing a pathological condition, e.g. for modulating gene expression in gastrointestinal inflammation, or for treating cancers or genetic disorders.

PS Claim 1; Page 84; 108pp; English.

CC The present sequence is one of a number of isolated human polynucleotides which are useful in gene therapy, and for diagnosing a pathological condition or a susceptibility to it. In particular, the polynucleotides are useful for modulating gene expression in gastrointestinal

CC inflammation. The polynucleotides are useful for chromosome  
CC identification, controlling gene expression through triple helix  
CC formation or antisense DNA or RNA, or identifying individuals from minute  
CC biological samples using DNA-based identification techniques. The  
CC polynucleotides can also be used as an alternative to restriction  
CC fragment length polymorphism (RFLP), by determining the actual base-by-  
CC base DNA sequences of selected portions of an individual's genome. The  
CC polynucleotides may also be used as molecular weight markers on Southern  
CC gels, as diagnostic probes for the presence of a specific mRNA, as a  
CC probe to substract-out known sequences in the process of discovering novel  
CC polynucleotides, or as an antigen to elicit an immune response. The  
CC polypeptides are useful in diagnostic procedures to detect a disorder.  
CC The polynucleotides and polypeptides are useful for preventing, treating  
CC or ameliorating immune system disorders, genetic disorders, cancers, some  
CC autoimmune disorders, or infections. The polynucleotides and polypeptides  
CC are also useful for differentiating, proliferating or attracting cells,  
CC leading to the regeneration of tissues, especially in wounds or burns.  
CC The polypeptides and polynucleotides may also be used as a food additive  
CC or preservative

XX Sequence 294 BP; 106 A; 52 C; 41 G; 95 T; 0 U; 0 Other;

Query Match 100.0%; Score 294; DB 4; Length 294;  
Best Local Similarity 100.0%; Pred. No. 1.2e-56;  
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 CGGCCTTAAGGTTCTCTGACAGTTGTCTCAAGAAGATCAGGATTGAGCATCTCTTACT 60  
Db 1 CGGCCTTAAGGTTCTCTGACAGTTGTCTCAAGAAGATCAGGATTGAGCATCTCTTACT 60  
Qy 61 TCAATATTCTTCCAGTGGTTTTGTTATCTCTCAATGTTATAGTGAATATAATAATA 120  
Db 61 TCAATATTCTTCCAGTGGTTTTGTTATCTCTCAATGTTATAGTGAATATAATAATA 120  
Qy 121 CATGAGAAAAAGCATAAATCAACTATTGAATAACGAAATGTGAACCTCTATGTAAT 180  
Db 121 CATGAGAAAAAGCATAAATCAACTATTGAATAACGAAATGTGAACCTCTATGTAAT 180  
Qy 181 GCAACTCAGGCTGCTACCGTTTTTAATTTCTCTTAATCTTTAAAAACAATACTATCT 240  
Db 181 GCAACTCAGGCTGCTACCGTTTTTAATTTCTCTTAATCTTTAAAAACAATACTATCT 240  
Qy 241 TTACTAAAGATGTTTCAGAGTAAGTAAATAACAAAGAGAGTAAATTTCCCAAAA 294  
Db 241 TTACTAAAGATGTTTCAGAGTAAGTAAATAACAAAGAGAGTAAATTTCCCAAAA 294

RESULT 2  
ABD33179  
ID ABD33179 standard; DNA; 100445 BP.

XX ABD33179;  
XX  
XX 18-NOV-2004 (first entry)  
XX Human cancer-associated (CA) gene HD07-026.

XX Human; cancer-associated protein; CAP; cancer-associated gene; CA; gene;  
KW ds; cancer; cytostatic.  
XX Homo sapiens.  
XX WO2004058146-A2.  
XX  
XX 15-JUL-2004.  
XX  
XX 15-DEC-2003; 2003WO-US040081.  
XX  
XX 17-DEC-2002; 2002US-00322281.  
XX (SAGR-) SAGRES DISCOVERY INC.  
XX  
XX Morris DW, Malandro MS;

XX

DR WPI; 2004-499109/47.

XX Novel human cancer associated protein encoded within open reading frame  
XX of cancer associated gene, useful as targets for diagnosing cancer.

XX Claim 16; SEQ ID NO 170; 182pp; English.

XX The invention relates to cancer-associated proteins (CAP) and the cancer-  
XX associated (CA) nucleic acids encoding them. The invention also relates  
XX to a method for treating cancers involving administering to a patient an  
XX inhibitor of CAP, and a method of screening for anticancer activity in a  
XX potential drug involving providing a cell that expresses a CA gene,  
XX contacting a tissue sample derived from a cancer cell with an anticancer  
XX drug candidate and monitoring the effect of the anticancer drug candidate  
XX on expression of the CA gene. The CAP proteins are useful for detecting  
XX cancer associated with expression of a CAP protein in a test cell sample  
XX and for screening for a bioactive agent capable of modulating the  
XX activity of a CAP protein. The CA nucleic acids are useful for diagnosing  
XX cancer, involving determining the expression of a CA nucleic acid in a  
XX tissue. This sequence represents a human CA gene of the invention. Note:  
XX The sequence data for this patent did not form part of the printed  
XX specification, but was obtained in electronic format directly from WIPO  
XX at ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 100445 BP; 25647 A; 23783 C; 23245 G; 27750 T; 0 U; 20 Other;

Query Match 98.0%; Score 288.2; DB 13; Length 100445;  
Best Local Similarity 99.0%; Pred. No. 5.1e-55;  
Matches 290; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
Qy 1 CGGCCTTAAGGTTCTCTGACAGTTGTCTCAAGAAGATCAGGATTGAGCATCTCTTACT 60  
Db 61160 CGGCCTTAAGGTTCTCTGACAGTTGTCTCAAGAAGATCAGGATTGAGCATCTCTTACT 61219  
Qy 61 TCAATATTCTTCCAGTGGTTTTGTTATCTCTCAATGTTATAGTGAATATAATAATA 120  
Db 61220 TCAATATTCTTCCAGTGGTTTTGTTATCTCTCAATGTTATAGTGAATATAATAATA 61279  
Qy 121 CATGAGAAAAAGCATAAATCAACTATTGAATAACGAAATGTGAACCTCTATGTAAT 180  
Db 61280 CATGAGAAAAAGCATAAATCAACTATTGAATAACGAAATGTGAACCTCTATGTAAT 61339  
Qy 181 GCAACTCAGGCTGCTACCGTTTTTAATTTCTCTTAATCTTTAAAAACAATACTATCT 240  
Db 61340 GCAACTCAGGCTGCTACCGTTTTTAATTTCTCTTAATCTTTAAAAACAATACTATCT 61399  
Qy 241 TTACTAAAGATGTTTCAGAGTAAGTAAATAACAAAGAGAGTAAATTTCCCAAAA 293  
Db 61400 TTACTAAAGATGTTTCAGAGTAAGTAAATAACAAAGAGAGTAAATTTCCCAAAA 61452

RESULT 3  
ABL33173/c  
ID ABL33173 standard; DNA; 15714 BP.

XX ABL33173;  
XX

XX 26-MAR-2002 (first entry)  
XX Human immune system associated gene SEQ ID NO: 1146.

XX Human; immune system disease; cytosine methylation; antiasthmatic;  
KW antiarteriosclerotic; antianemic; cytostatic; nootropic;  
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;  
KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;  
KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;  
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;  
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;  
KW ds.  
XX Homo sapiens.  
XX

```
PN WO200200928-A2.
XX
PD 03-JAN-2002.
XX
PF 02-JUL-2001; 2001WO-EP007537.
XX
PR 30-JUN-2000; 2000DE-01032529.
XX
PR 01-SEP-2000; 2000DE-01043826.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2002-130909/17.
XX
XX Nucleic acid comprising fragment of chemically modified gene, useful for
PT diagnosis and treatment of diseases associated with abnormal cytosine
PT methylation.
XX
PS Claim 1; SEQ ID NO 1146; 32pp + Sequence Listing; German.
XX
CC The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention
XX
SQ Sequence 15714 BP; 4163 A; 246 C; 3762 G; 7543 T; 0 U; 0 Other;
Query Match 16.1%; Score 47.2; DB 6; Length 15714;
Best Local Similarity 52.9%; Pred. No. 0.52;
Matches 129; Conservative 0; Mismatches 108; Indels 7; Gaps 1;
QY 31 AAGAAGATCAGAGATTGAGCATCTCTTACTTCAAAATTATTTCTCCAGTGGTTTGTATC 90
Db 6545 ACGAAATCCAAACTCGCTACTTTATTTCTCCTCTTATTTCTTCTCATCCCAATC 6486
QY 91 TTCTCAATGTTTACTGAAATATATATACATGAGAAAAGCATAACTCACTATTG 150
Db 6485 CTTTCAACTTAAATATATATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAA 6426
QY 151 AATAACGAAAATGTGAACCTCTATGTAACCTCAACTCAGGCTGCCTACCGTTTAAATTT 210
Db 6425 AAAATCCCAAAAAAATATATATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAA 6373
QY 211 CTCTCTAATCTTTAAAAACAATACTATCTTTTCTAAAGATGTTTCAGAGTAAGATAAA 270
Db 6372 TTCCCTCATAATTAATAACAATACCATATTACCTAAATTTTAAATTCACAATAACATCAA 6313
QY 271 TACA 274
Db 6312 AAAA 6309
RESULT 4
ABQ67058/c
ID ABQ67058 standard; DNA; 15714 BP.
XX
AC ABQ67058;
XX
DT 28-AUG-2002 (first entry)
XX
DE Human angiogenesis associated polynucleotide SEQ ID NO 88.
XX
KW Human; angiogenesis; methylation; eye disease; glaucoma; tumour;
KW inflammation; rheumatoid arthritis; diabetic retinopathy; antiulcers;
KW macular degeneration; inflammatory bowel disease; Crohn's disease;
KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW antiarteriosclerotic; ds.
```

```
OS Homo sapiens.
XX
PN WO200246454-A2.
XX
PD 13-JUN-2002.
XX
PF 06-DEC-2001; 2001WO-EP014320.
XX
PR 06-DEC-2000; 2000DE-01061338.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Schacht O;
XX
DR WPI; 2002-500450/53.
XX
XX New nucleic acid fragments from chemically treated angiogenesis-
PT associated genes, useful for determining methylation status, e.g. in
PT diagnosis or treatment of cancer.
XX
PS Claim 1; SEQ ID NO 88; 4lpp + Sequence Listing; German.
XX
CC The invention relates to a nucleic acid (I) comprising a segment of 18
CC bases of chemically pretreated DNA of angiogenesis-associated genes (II)
CC having sequences (ABQ66971-ABQ67178) or their complements. (I), also
CC related oligomers, are used to evaluate the methylation status and/or
CC single-nucleotide polymorphisms, in angiogenesis-related genes, for
CC diagnosis and treatment of eye diseases, proliferative retinopathy,
CC neovascular glaucoma, solid tumours, inflammation, rheumatoid arthritis,
CC diabetic retinopathy, macular degeneration caused by neovascularisation,
CC psoriasis, arteriosclerosis, inflammatory bowel diseases, ulcers and
CC Crohn's disease. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 15714 BP; 4165 A; 246 C; 3759 G; 7544 T; 0 U; 0 Other;
Query Match 16.1%; Score 47.2; DB 6; Length 15714;
Best Local Similarity 52.9%; Pred. No. 0.52;
Matches 129; Conservative 0; Mismatches 108; Indels 7; Gaps 1;
QY 31 AAGAAGATCAGAGATTGAGCATCTCTTACTTCAAAATTATTTCTCCAGTGGTTTGTATC 90
Db 6545 ACGAAATCCAAACTCGCTACTTTATTTCTCCTCTTATTTCTTCTCATCCCAATC 6486
QY 91 TTCTCAATGTTTACTGAAATATATATAATACATGAGAAAAGCATAACTCACTATTG 150
Db 6485 CTTTCAACTTAAATATATATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAA 6426
QY 151 AATAACGAAAATGTGAACCTCTATGTAACCTCAACTCAGGCTGCCTACCGTTTAAATTT 210
Db 6425 AAAATCCCAAAAAAATATATATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAA 6373
QY 211 CTCTCTAATCTTTAAAAACAATACTATCTTTTCTAAAGATGTTTCAGAGTAAGATAAA 270
Db 6372 TTCCCTCATAATTAATAACAATACCATATTACCTAAATTTTAAATTCACAATAACATCAA 6313
QY 271 TACA 274
Db 6312 AAAA 6309
RESULT 5
ABL32224/c
ID ABL32224 standard; DNA; 14924 BP.
XX
AC ABL32224;
XX
DT 26-MAR-2002 (first entry)
XX
DE Human immune system associated gene SEQ ID NO: 197.
XX
KW Human; immune system disease; cytosine methylation; antiasthmatic;
```

KW antiarteriosclerotic; antianemic; cytotetic; nootropic;  
 KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;  
 KW antirheumatic; antithyroid; antidiabetic; antipeptidic;  
 KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;  
 KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;  
 KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;  
 KW ds.  
 XX Homo sapiens.  
 OS  
 XX WO200200928-A2.  
 PN  
 XX 03-JAN-2002.  
 PD  
 XX  
 XX 02-JUL-2001; 2001WO-EP007537.  
 FF  
 XX 30-JUN-2000; 2000DE-01032529.  
 PR  
 XX 01-SEP-2000; 2000DE-01043826.  
 PR  
 XX (EPIG-) EPIGENOMICS AG.  
 PA  
 XX Olek A, Piepenbrock C, Berlin K;  
 PI  
 XX WPI; 2002-130909/17.  
 DR  
 XX Nucleic acid comprising fragment of chemically modified gene, useful for  
 PT diagnosis and treatment of diseases associated with abnormal cytosine  
 PT methylation.  
 PT  
 XX Claim 1; SEQ ID NO 197; 32pp + Sequence Listing; German.  
 PS  
 XX  
 XX The present invention provides a number of human immune system associated  
 CC genes which are modified by the methylation of cytosines. The sequences  
 CC can be used in the diagnosis and treatment of immune system disorders,  
 CC including eye diseases such as retinopathy, neovascular glaucoma and  
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid  
 CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,  
 CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel  
 CC diseases. The present sequence is a gene of the invention  
 CC  
 XX  
 SQ Sequence 14924 BP; 4130 A; 179 C; 3007 G; 7602 T; 0 U; 6 Other;  
 Query Match 15.2%; Score 44.8; DB 6; Length 14924;  
 Best Local Similarity 48.1%; Pred. No. 1.8;  
 Matches 127; Conservative 0; Mismatches 137; Indels 0; Gaps 0;  
 QY 31 AAGAAGATCAGGATTTGAGCATCTCTTACTTCAAAATATTCTCCAGTGGTTTTGTTATC 90  
 DB 12072 AATAAATAAATAACTTTTCTCTTCAAAAATATTATTTCAAAATATATTTTCTTTAA 12013  
 QY 91 TTCTCAATGTTATAGTGAATATATAATATACATGCAGAAAAAGCATAAATCAACTATTG 150  
 DB 12012 TAATACTATTATTATCTTTAAACCATTTATATAAATAATATAAACAATAATTCAT 11953  
 QY 151 AATAACGAAATGTGAACCTCTATGTAACCTCAGGCTGCCTACCGTTTTTAATTT 210  
 DB 11952 ATCTATCAATAATATAATAATAATAATAATAATAATAATAATAATAATAATAATA 11893  
 QY 211 CTCTCTAATCTTTAAAAACAATACTATCTTTACTAAAGATGTTTCAGAGTAAGATAAAA 270  
 DB 11892 ATAATAATTTTAAATTTCACTACCTTAATTAACAAAAAACCTTAAACATAATAATA 11833  
 QY 271 TACAAAGAGAGTAATTCCTCAAAAA 294  
 DB 11832 TACAATCAATAATTCCTCAAAAA 11809  
 RESULT 6  
 ABL54321/c  
 ID ABL54321 standard; DNA; 14924 BP.  
 XX  
 AC ABL54321;  
 XX

DT 29-JUL-2002 (first entry)  
 DE Chemically treated apoptosis gene #11.  
 XX  
 KW Apoptosis; HIV; Bloom syndrome; cardiopathy; neurodegenerative disorder;  
 KW Herpes simplex virus; renal ischaemia; amyotrophic lateral sclerosis;  
 KW cancer; ds.  
 XX  
 OS Unidentified.  
 XX  
 XX WO200177164-A2.  
 PN  
 XX 18-OCT-2001.  
 PD  
 XX  
 XX 06-APR-2001; 2001WO-EP003969.  
 FF  
 XX 06-APR-2000; 2000DE-01019058.  
 PR  
 XX 07-APR-2000; 2000DE-01019173.  
 PR  
 XX 30-JUN-2000; 2000DE-01032529.  
 PR  
 XX 01-SEP-2000; 2000DE-01043826.  
 PR  
 XX (EPIG-) EPIGENOMICS AG.  
 PA  
 XX Olek A, Piepenbrock C, Berlin K;  
 PI  
 XX WPI; 2002-017444/02.  
 DR  
 XX Chemically modified sequences of genes associated with apoptosis are  
 PT useful to determine methylation patterns of genomic DNA samples for  
 PT diagnosis of associated diseases such as cancer.  
 PT  
 XX Claim 1; Seq ID #21; 24pp; English.  
 PS  
 XX This invention relates to chemically pre-treated DNA of genes associated  
 CC with apoptosis. The nucleic acids are used to allocate patients for  
 CC specific therapy for HIV infection, Bloom syndrome, cardiopathy, aging,  
 CC neurodegenerative disorders, Herpes simplex virus infection, renal  
 CC ischaemia, amyotrophic lateral sclerosis, solid tumours and cancers. This  
 CC nucleotide sequence represents a chemically treated apoptosis gene. Even  
 CC SEQ ID numbers are the complementary DNA strands to the odd SEQ ID  
 CC numbers. The sequence data for this patent is not represented in the  
 CC printed specification but is based on information supplied by the  
 CC European patent office  
 CC  
 SQ Sequence 14924 BP; 4130 A; 179 C; 3007 G; 7602 T; 0 U; 6 Other;  
 Query Match 15.2%; Score 44.8; DB 6; Length 14924;  
 Best Local Similarity 48.1%; Pred. No. 1.8;  
 Matches 127; Conservative 0; Mismatches 137; Indels 0; Gaps 0;  
 QY 31 AAGAAGATCAGGATTTGAGCATCTCTTACTTCAAAATATTCTCCAGTGGTTTTGTTATC 90  
 DB 12072 AATAAATAAATAACTTTTCTCTTCAAAAATATTATTTCAAAATATATTTTCTTTAA 12013  
 QY 91 TTCTCAATGTTATAGTGAATATATAATATACATGCAGAAAAAGCATAAATCAACTATTG 150  
 DB 12012 TAATACTATTATTATCTTTAAACCATTTATATAAATAATATAAACAATAATTCAT 11953  
 QY 151 AATAACGAAATGTGAACCTCTATGTAACCTCAGGCTGCCTACCGTTTTTAATTT 210  
 DB 11952 ATCTATCAATAATATAATAATAATAATAATAATAATAATAATAATAATAATAATA 11893  
 QY 211 CTCTCTAATCTTTAAAAACAATACTATCTTTACTAAAGATGTTTCAGAGTAAGATAAAA 270  
 DB 11892 ATAATAATTTTAAATTTCACTACCTTAATTAACAAAAAACCTTAAACATAATAATA 11833  
 QY 271 TACAAAGAGAGTAATTCCTCAAAAA 294  
 DB 11832 TACAATCAATAATTCCTCAAAAA 11809  
 RESULT 7  
 ABL33019/c

ID	ABL33019 standard; DNA; 5311 BP.
XX	ABL33019;
XX	26-MAR-2002 (first entry)
XX	Human immune system associated gene SEQ ID NO: 992.
XX	Human; immune system disease; cytosine methylation; antiasthmatic;
KW	antiarteriosclerotic; antianaemic; cytotatic; neutrotropic;
KW	neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW	antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW	antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW	acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW	neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
KW	ds.
OS	Homo sapiens.
XX	WO200200928-A2.
XX	03-JAN-2002.
XX	02-JUL-2001; 2001WO-EP007537.
XX	30-JUN-2000; 2000DE-01032529.
XX	01-SEP-2000; 2000DE-01043826.
XX	(EPIG-) EPIGENOMICS AG.
PA	Olek A, Piepenbrock C, Berlin K;
PI	WPI; 2002-130909/17.
XX	Nucleic acid comprising fragment of chemically modified gene, useful for
PT	diagnosis and treatment of diseases associated with abnormal cytosine
PT	methylation.
XX	Claim 1; SEQ ID NO 992; 32pp + Sequence Listing; German.
XX	The present invention provides a number of human immune system associated
CC	genes which are modified by the methylation of cytosines. The sequences
CC	can be used in the diagnosis and treatment of immune system disorders,
CC	including eye diseases such as retinopathy, neovascular glaucoma and
CC	macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC	leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC	rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC	diseases. The present sequence is a gene of the invention
XX	Sequence 5311 BP; 1696 A; 41 C; 863 G; 2711 T; 0 U; 0 Other;
XX	Query Match 15.0%; Score 44; DB 6; Length 5311;
XX	Best Local Similarity 53.0%; Pred. No. 2.4;
XX	Matches 104; Conservative 0; Mismatches 100; Indels 0; Gaps 0;
Qy	91 TTCTCAATGTTATAGTCAAAATATATATACATGCAGAAAAAGCATATAATCACTATTG 150
Db	2082 TACTTACAATACATTATTAACAACCCATTAACTAACAAATAAAAAAATAATTAATA 2033
Qy	151 AATAACGAAATGTGAACCTCTATGTAACTGCAACTCAGGCTGCCTACCGTTTTTAATTT 210
Db	2022 AAAAAAANAATTTATTAAACCTCAATATAAAAAAATAAACATTTCTCTCTAATTTTATAAA 1963
Qy	211 CTCTCTAATCTTTAAAAACAATACTATCTTTACTAAAGATGTTTCAGAGTAAAGATAAA 270
Db	1962 ACGCAAAATTTTAAAAATAACATAAATTAATTTCACTTAATTTACTAATTTAAAAAANA 1903
Qy	271 TACAAAGAGAGTAATTTCCAAAAA 294
Db	1902 TTTAAACACTAAATTTCCAAAAA 1879
XX	RESULT 8

ABN80257/C	ID	ABN80257 standard; DNA; 6197 BP.
XX	AC	XX
XX	ABN80257;	XX
XX	XX	XX
DT	15-JUL-2002 (first entry)	XX
XX	Human chemically modified disease associated gene SEQ ID NO 274.	XX
DE	Human; development; homeobox gene; HOX; diabetes; cancer; apoptosis;	XX
KW	heart disease; epilepsy; histone deacetylation; muscular dystrophy;	XX
KW	dwarfism; single nucleotide polymorphism; SNP; cytosine methylation;	XX
KW	antidiabetic; cytostatic; anticonvulsant; ds.	XX
XX	Homo sapiens.	OS
OS	Synthetic.	OS
XX	WO200200927-A2.	XX
PN	03-JAN-2002.	XX
PD	02-JUL-2001; 2001WO-EP007536.	XX
XX	30-JUN-2000; 2000DE-01032529.	XX
PR	01-SEP-2000; 2000DE-01043826.	XX
PR	(EPIG-) EPIGENOMICS AG.	XX
XX	Olek A, Piepenbrock C, Berlin K;	XX
PI	WPI; 2002-130908/17.	XX
XX	Novel nucleic acid useful for diagnosis and therapy of diseases	XX
DR	associated with development genes such as diabetes, comprises a sequence	XX
XX	of a segment of chemically pretreated DNA of genes associated with	XX
PT	development.	XX
PT	Claim 1; SEQ ID NO 274; 27pp; English.	XX
PS	The invention relates to a nucleic acid (I) comprising a sequence at	XX
XX	least 18 bases in length of a segment of chemically pretreated DNA (II)	CC
CC	of genes associated with development selected from 87 genes listed in the	CC
CC	specification such as ACCPN, ADFN, or AFDI and comprising one of 350	CC
CC	sequences (ABN79984-ABN80333) or their complements. The invention is	CC
CC	useful for the diagnosis or therapy of diseases associated with	CC
CC	development genes, in particular disease related to homeobox containing	CC
CC	genes (HOX), like diabetes, cancer, apoptosis related diseases, syndromes	CC
CC	associated with congenital heart disease, epilepsy, diseases related to	CC
CC	histone deacetylation, Curarino syndrome, diseases related with the	CC
CC	development of the brain and limb girdle muscular dystrophy and dwarfism.	CC
CC	Oligomers specific to each of the genes are useful for detecting the	CC
CC	methylation state of all CpG dinucleotides within the 350 sequences or	CC
CC	(II) and their complementary sequences, as primer oligonucleotides for	CC
CC	the amplification of the 350 sequences, (II) and/or their complements and	CC
CC	as oligomer probes for detecting the cytosine methylation state and/or	CC
CC	single nucleotide polymorphisms (SNPs). Note: The sequence data for this	CC
CC	patent did not form part of the printed specification but is based on	CC
CC	sequence information supplied to Derwent by the European Patent Office	CC
XX	Sequence 6197 BP; 1637 A; 45 C; 1206 G; 3309 T; 0 U; 0 Other;	XX
SQ	Query Match 14.9%; Score 43.8; DB 6; Length 6197;	
	Best Local Similarity 52.5%; Pred. No. 2.7;	
	Matches 96; Conservative 0; Mismatches 87; Indels 0; Gaps 0;	
QY	108 AAATATATATATATACATGCAGAAAAAAGCATTAATCAACTATTGTAATAACGAAATGTCAA 167	
DB	2581 AAAAATATATACGATCTAAATATATAAAATATAAAATATATAAAATATAAAATATAAAAC 2522	
QY	168 CCTCTATGTAACTGCAACTCAGGCTCGCTACCGGTTTTTAATTTCTCTTAATCTTTAAA 227	
DB	2521 CTATATTTCAACAACAACAACAAAATAACTATCGTCAATAATAATTTCAATTATACATATA 2462	





CC alternative splicing events, in detecting and characterising gross  
CC alterations in the genomic locus that includes their exon, in assessing  
CC smaller genomic alterations, in priming the synthesis of nucleic acids,  
CC or in expressing the ORF-encoded peptide. The present sequence is a human  
CC single exon probe of the invention. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from USPTO at  
CC seqdata.uspto.gov/sequence.html?DocID=20030194704  
XX

Query Match 14.8%; Score 43.6; DB 12; Length 510;  
Best Local Similarity 48.8%; Pred. No. 2.2;  
Matches 118; Conservative 0; Mismatches 124; Indels 0; Gaps 0;  
SQ Sequence 510 BP; 158 A; 79 C; 69 G; 204 T; 0 U; 0 Other;

QY 2 GGCTTAAAGTTCCTCTGACAGTGTCTCAAGAGATCAGGATTTGAGCATCTCTTACTT 61  
DB 225 GGCACTAAGTCTCTCTGAGTGTCTCTTTCAGTCTTACTTATTTAGATCTTCTT 284  
QY 62 CAAATTTATTTCTCCAGTGGTTTGTGTTATCTTCTCAATGTTATAGTGAATATAATATAC 121  
DB 285 CAGACTCTGAATTAAGTCTCTGTTGTAGTCAAGATGATTTGATCTTTTAAATAATGGA 344  
QY 122 ATGCAGAAAAGCAATAATCACTATTGAAATACGAAATGTGAACCTCTATGTAATG 181  
DB 345 GGAAGAAGAAGATTAATAATATATGAGACATTTTTCAAAAGTTTGAAGTCCAAATGTAAT 404  
QY 182 CAACTCAGGTGCTACCGTTTTTAAATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 241  
DB 405 TATGTTATCTGCTTTTAAATCACTTTTTTTTTTGAAATCTTAAATAATCCCATACTT 464  
QY 242 TA 243  
DB 465 CA 466

RESULT 11  
ID ABL32901/c  
ID ABL32901 standard; DNA; 5204 BP.  
XX  
AC ABL32901;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Human immune system associated gene SEQ ID NO: 874.  
XX  
KW Human; immune system disease; cytosine methylation; antiasthmatic;  
KW antiarteriosclerotic; anti-naemic; cytosine; neurotropic;  
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;  
KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;  
KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;  
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;  
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;  
KW db.  
XX  
OS Homo sapiens.  
XX  
PN WO200200928-A2.  
XX  
PD 03-JAN-2002.  
XX  
PF 02-JUL-2001; 2001WO-EP007537.  
XX  
PR 30-JUN-2000; 2000DE-01032529.  
PR 01-SEP-2000; 2000DE-01043826.  
XX  
PA (EPIG-) EPIGENOMICS AG.  
XX  
PI Olek A, Piepenbrock C, Berlin K;  
XX  
XX WPI; 2002-130909/17.  
XX  
PT Nucleic acid comprising fragment of chemically modified gene, useful for

PT diagnosis and treatment of diseases associated with abnormal cytosine  
PT methylation.  
XX  
PS Claim 1; SEQ ID NO 874; 32pp + Sequence Listing; German.

XX The present invention provides a number of human immune system associated  
CC genes which are modified by the methylation of cytosines. The sequences  
CC can be used in the diagnosis and treatment of immune system disorders,  
CC including eye diseases such as retinopathy, neovascular glaucoma and  
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid  
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,  
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel  
CC diseases. The present sequence is a gene of the invention  
XX  
SQ Sequence 5204 BP; 1463 A; 72 C; 1014 G; 2655 T; 0 U; 0 Other;

Query Match 14.8%; Score 43.4; DB 6; Length 5204;  
Best Local Similarity 50.2%; Pred. No. 3.2;  
Matches 107; Conservative 0; Mismatches 106; Indels 0; Gaps 0;  
QY 82 TTTGTTATCTTCTCAATGTTATAGTGAATATAATATACATGCAGAAAAAGCAATAAT 141  
DB 2290 TTTTTCACGCTAAACGCCCTAATAACAAAAAATAAAAAAATAAAAAA 2231  
QY 142 CAACTATTGAAATACGAAATGTGAACCTCTATGTAACTGCAACTCAGCTCGCTACCGT 201  
DB 2230 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 2171  
QY 202 TTTTAAATTTCTCTCTAAATCTTTAAAAACAATACTATCTTTTAAAGATGTTTCAGCT 261  
DB 2170 TAAATTCATTTCT 2111  
QY 262 AAGATAAAATACAAAGAGAGATTAATTCCTCAAAA 294  
DB 2110 AAAAAAATAATACTAATAATAACCCCAATA 2078

RESULT 12  
ID ABK28258/c  
ID ABK28258 standard; DNA; 8170 BP.  
XX  
AC ABK28258;  
XX  
DT 23-APR-2002 (first entry)  
XX  
DE DNA transcription associated complementary genomic DNA #66.  
XX  
KW DNA transcription associated gene; peptide nucleic acid; PNA-oligomer;  
KW PNA; cytosine methylation state; SNP; retroviral infection; gene; db;  
KW single nucleotide polymorphism; adenosine deaminase deficiency; cancer;  
KW viral infection; Sezary syndrome; haematological disorder; tuberculosis;  
KW immunological disorder; Werner syndrome; developmental disorder;  
KW psoriasis; Rieger's syndrome; neurological disorder; erythrocytosis;  
KW neurodegenerative disorder; Waardenburg syndrome; Niemann-Pick disease;  
KW myelodysplastic syndrome; myocardial infarction; hypertension; arthritis;  
KW angioedema; congenital heart disease; HDR syndrome; gene therapy;  
KW polyglutamine disorder; solid tumour.  
XX  
OS Unidentified.  
XX  
PN WO200192565-A2.  
XX  
PD 06-DEC-2001.  
XX  
PF 06-APR-2001; 2001WO-EP003973.  
XX  
PR 06-APR-2000; 2000DE-01019058.  
PR 07-APR-2000; 2000DE-01019173.  
PR 30-JUN-2000; 2000DE-01032529.  
PR 01-SEP-2000; 2000DE-01043826.  
XX  
XX (EPIG-) EPIGENOMICS AG.  
XX  
XX

PI Olek A, Piepenbrock C, Berlin K;  
 XX WPI; 2002-090046/12.  
 XX New nucleic acids or oligomers, useful for diagnosing or treating  
 PT diseases associated with DNA transcription, e.g. immunological disorders,  
 PT Werner syndrome, psoriasis, myocardial infarction, solid tumors or  
 PT cancer.  
 XX  
 XX Claim 1; SEQ ID NO 132; 32pp; English.  
 XX  
 CC The invention relates to a nucleic acid, which comprises a segment of the  
 CC chemically pretreated DNA of genes associated with DNA transcription from  
 CC one of 346 sequences, and an oligomer, in particular an oligonucleotide  
 CC or peptide nucleic acid (PNA)-oligomer that hybridises to or is identical  
 CC to the chemically pretreated DNA of genes associated with DNA  
 CC transcription. The set of oligomer probes are useful for detecting the  
 CC cytosine methylation state and/or single nucleotide polymorphisms (SNPs)  
 CC in a chemically pretreated genomic DNA. The nucleic acids are useful for  
 CC diagnosing or treating diseases associated with DNA transcription  
 CC (particularly with the methylation status), e.g. adenosine deaminase  
 CC deficiency, viral infection, retroviral infection, Sezary syndrome,  
 CC haematological disorders, immunological disorders, Werner syndrome,  
 CC tuberculosis, developmental disorders, psoriasis, Rieger's syndrome,  
 CC neurological disorders, neurodegenerative disorders, Waardenburg  
 CC syndrome, Niemann-Pick disease, myelodysplastic syndrome, myocardial  
 CC infarction, hypertension, angiogenesis, erythropoiesis, congenital heart  
 CC disease, HDR syndrome, arthritis, polyglutamine disorders, solid tumours  
 CC or cancer. Sequences ABK28127-ABK28472 represent DNA transcription  
 CC associated genomic DNA molecules of the invention. Note: The sequence  
 CC data for this patent did not form part of the printed specification but  
 CC was obtained in electronic format directly from the European Patent  
 CC Office  
 XX  
 SQ Sequence 8170 BP; 2790 A; 69 C; 1334 G; 3977 T; 0 U; 0 Other;  
 Query Match 14.8%; Score 43.4; DB 6; Length 8170;  
 Best Local Similarity 50.2%; Pred. No. 3.4;  
 Matches 107; Conservative 0; Mismatches 106; Indels 0; Gaps 0;  
 QY 82 TTGTGTTCTCTCAATGTTATAGTGAATATATAATACATGAGAAAAAGCATAAAT 141  
 DB 2628 TTTTAAATCTAAAACCTTTTATATAAATATTTATATAATATAAACAATATAATA 2569  
 QY 142 CAACATTATTGAATAACGAAATGTGAACCTCTATGTAACCTGCAACTGCGCTACCGT 201  
 DB 2568 ATATTATAAATAACAATATTAAAAAATACTACTTCTTTATCTTAAAAATTTATTCGAT 2509  
 QY 202 TTTTAAATTTCTCTCAATCTTTTAAAAACAATACTATCTTTTACTAAAGATGTTTCAGAGT 261  
 DB 2508 TATTAAATTTACATTTTACTATCTTAAAAACAACCTATATAAACCCTTATATAAACAACAAC 2449  
 QY 262 AAGATAAAATACAAAGAGAGTAATTCACAAAAA 294  
 DB 2448 AATCTAAATATCAACAACAATAAATTTTCATATA 2416  
 RESULT 13  
 ABL33407/c  
 ID ABL33407 standard; DNA; 8423 BP.  
 XX  
 AC ABL33407;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE Human immune system associated gene SEQ ID NO: 1380.  
 XX  
 KW Human; immune system disease; cytosine methylation; antiasthmatic;  
 KW antiarteriosclerotic; antianaemic; cytostatic; nootropic;  
 KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;  
 KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;  
 KW antinflammatory; cancer; eye disease; arteriosclerosis; anaemia;  
 KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;  
 KW  
 KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;  
 ds.  
 KW  
 XX Homo sapiens.  
 XX WO200200928-A2.  
 XX  
 XX 03-JAN-2002.  
 XX  
 XX 02-JUL-2001; 2001WO-EP007537.  
 XX  
 XX 30-JUN-2000; 2000DE-01032529.  
 XX  
 XX 01-SEP-2000; 2000DE-01043826.  
 XX  
 XX (EPIG-) EPIGENOMICS AG.  
 XX  
 XX Olek A, Piepenbrock C, Berlin K;  
 XX WPI; 2002-130909/17.  
 XX  
 XX Nucleic acid comprising fragment of chemically modified gene, useful for  
 PT diagnosis and treatment of diseases associated with abnormal cytosine  
 PT methylation.  
 XX  
 XX Claim 1; SEQ ID NO 1380; 32pp + Sequence Listing; German.  
 XX  
 CC The present invention provides a number of human immune system associated  
 CC genes which are modified by the methylation of cytosines. The sequences  
 CC can be used in the diagnosis and treatment of immune system disorders,  
 CC including eye diseases such as retinopathy, neovascular glaucoma and  
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid  
 CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,  
 CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel  
 CC diseases. The present sequence is a gene of the invention  
 XX  
 SQ Sequence 8423 BP; 2529 A; 214 C; 1834 G; 3845 T; 0 U; 1 Other;  
 Query Match 14.8%; Score 43.4; DB 6; Length 8423;  
 Best Local Similarity 51.9%; Pred. No. 3.4;  
 Matches 98; Conservative 0; Mismatches 91; Indels 0; Gaps 0;  
 QY 86 TTATCTTCTCAATGTTATAGTGAATATATAATACATGCAAAAAAGCAATAAATCAAC 145  
 DB 1174 TCATCCCAAAACATAACCTCCCAATATATCTCTCAATATATAAACAACCTAAATAAAC 1115  
 QY 146 TATTGAATAACGAAATGTGAACCTCTATGTAACCTGCAACTGCGCTACCGTTTTT 205  
 DB 1114 AAAAAACATTTTAAAAATAAAAACTTTAATTTAAATTAATAATCAATAAAAAAACTCTT 1055  
 QY 206 AATTTCTCTCAATCTTTAAAAACAATACTATCTTTTACTAAAGATGTTTCAGAGTAAAG 265  
 DB 1054 ACTTCTCCCAATATTTAAAAACAACAATAACGATATATAAAAAAATAAATAAATTTATA 995  
 QY 266 TAAAAATACA 274  
 DB 994 TAAATTACA 986  
 RESULT 14  
 ABL34584/c  
 ID ABL34584 standard; DNA; 14551 BP.  
 XX  
 AC ABL34584;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE Human metastasis associated gene SEQ ID NO: 137.  
 XX  
 KW Metastasis associated gene; cytostatic; gene therapy; cancer;  
 KW cytosine methylation; gene; ds.  
 XX  
 OS Homo sapiens.  
 XX

```
PN WO200177376-A2.
XX
XX
PD 18-OCT-2001.
XX
XX
PF 06-APR-2001; 2001WO-EP003970.
XX
XX
XX 06-APR-2000; 2000DE-01019058.
XX
XX 07-APR-2000; 2000DE-01019173.
PR
PR 30-JUN-2000; 2000DE-01032529.
PR
PR 01-SEP-2000; 2000DE-01043826.
XX
XX (EPIG-) EPIGENOMICS AG.
PA
XX
XX Olek A, Piepenbrock C, Berlin K;
PI
XX
XX WPI; 2002-010922/01.
DR
XX
XX New nucleic acid derived from chemically treated metastasis genes, useful
PT for diagnosis of cancers by analysis of cytosine methylation, also for
PT treatment.
XX
XX Claim 1; SEQ ID NO 137; 23pp + Sequence Listing; English.
PS
XX
XX The present invention provides a number of human metastasis associated
CC genes which are modified by cytosine methylation. The sequences can be
CC used in the diagnosis and treatment of cancer. The present sequence is
CC one of the genes of the invention
XX
XX Sequence 14551 BP; 4362 A; 107 C; 3194 G; 6888 T; 0 U; 0 Other;
SQ
Query Match 14.7%; Score 43.2; DB 6; Length 14551;
Best Local Similarity 52.7%; Pred. No. 4.1;
Matches 117; Conservative 0; Mismatches 103; Indels 2; Gaps 1;
OY 57 TACTTCAAATTTATCTTCCAGTGGTTTGTATCTTCTCAATGTTATAGTGAATATAAT 116
Db 1078 TAATACTAAACACTATATAAAAAATTTTCTCTATTTAATACATCTCAAACTAAATAAAAA 1019
OY 117 AATACATGCAGAAAAAGCATAAATCAACTATTGATTA--CGAAAAATGTGAACCTCTAT 174
Db 1018 CATACATTAATAATAAAAAATCAAAAAATTTAATAAAAAATTTCTAAATTTCTAAATAAAT 959
OY 175 GAACTGCAACTCAGGCTGCTTACCGTTTCTTAAATTTTCTCTAAATCTTTAAAAACAATA 234
Db 958 TCTATATAAAATCTCTCTACCCCAAAATCTTCTTCACTTAAACGTCGAAATATAAC 899
OY 235 CTATCTTTACTAAAGATGTTTCAGAGTAAGATAAAATACAA 276
Db 898 TTATTTAAATCAAATCTTTCAAATATAAATAATAAAAA 857
RESULT 15
ADS99845/c
ID ADS99845 standard; DNA; 14551 BP.
XX
XX ADS99845;
AC
XX
XX 02-DEC-2004 (first entry)
DT
XX
XX Bisulphite treated human gene associated with metastasis #69.
XX
XX Human; ds; gene; Bisulphite; metaerasis; cancer; cytostratic;
KW DNA methylation; matrix-assisted laser desorption/ionisation; MALDI;
KW electrospray; mass spectrometry; CpG dinucleotide; solid tumour.
XX
XX Homo sapiens.
OS
XX
XX US2003148327-A1.
PN
XX
XX 07-AUG-2003.
PD
XX
XX 21-JAN-2003; 2003US-00240485.
PF
XX
XX
```

```
PR 06-APR-2000; 2000DE-01019058.
PR
PR 07-APR-2000; 2000DE-01019173.
PR
PR 30-JUN-2000; 2000DE-01032529.
PR
PR 01-SEP-2000; 2000DE-01043826.
PR
PR 06-APR-2001; 2001WO-EP003970.
XX
XX (OLEK/) OLEK A.
PA
PA (PIEP/) PIEPENBROCK C.
PA
PA (BERL/) BERLIN K.
XX
XX Olek A, Piepenbrock C, Berlin K;
PI
XX
XX WPI; 2002-010922/01.
DR
XX
XX New nucleic acid derived from chemically treated metastasis genes, useful
PT for diagnosis of cancers by analysis of cytosine methylation, also for
PT treatment.
XX
XX Claim 1; SEQ ID NO 137; 9pp; English.
PS
XX
XX The invention relates to a nucleic acid comprising at least 18 bases from
CC a segment of the chemically pretreated DNA of genes associated with
CC metastasis, i.e. any of ADS9709-ADS9906 human genomic sequences or any
CC of the 19 sequences appearing as ADS9911-ADS9929. SEQ ID 2,4,6 etc are
CC the complements of SEQ ID 1,3,5, etc. Also included are an oligomer
CC (particularly an oligonucleotide or peptide nucleic acid) comprising at
CC least one base sequence of at least 9 bases which hybridises to (or is
CC identical with) the sequences referred to above, producing an array of
CC the oligomers on a carrier, obtaining genetic and/or epigenetic
CC parameters for diagnosis and/or therapy of diseases (or predisposition to
CC them) by analysis of cytosine methylation and a kit comprising a
CC bisulphite (disulphite or hydrogen sulphite) and the oligomers. In the
CC method of above 5-unmethylated cytosines in a genomic DNA sample are
CC converted chemically to uracil, or another base with hybridisation
CC properties different from those of cytosine, then fragments of the
CC treated DNA amplified (particularly by polymerase chain reaction) using
CC the oligomers and a polymerase (preferably heat stable) to produce
CC labelled amplicons. These are tested for hybridisation to an array of
CC oligomers and any hybridisation detected. The amplicons are labelled with
CC fluorescent or radioactive markers, or with a detachable mass marker to
CC allow their detection by mass spectrometry, specifically using the matrix
CC assisted laser desorption/ionisation (MALDI) or electrospray techniques.
CC To improve detection in the mass spectrometer, fragments formed in the
CC instrument have only a single net charge (positive or negative). The
CC genomic DNA is from e.g. a cell line, biopsy sample, blood, or paraffin-
CC embedded tissue sample. Oligonucleotides or peptide-nucleic acids that
CC are complementary to (or identical with) parts of the nucleic acids listed
CC above may be used as primers for amplification of the nucleic acids or
CC their complements, and for determining cytosine methylation status and/or
CC single nucleotide polymorphisms in metastasis-related genes. They can be
CC used for analysis of diseases associated with methylation of CpG
CC dinucleotides and to determine (epigenetic parameters for diagnosis
CC and/or therapy of disease (or predisposition)). The genomic DNA sequences
CC are useful for diagnosis and therapy of solid tumours and cancer. The
CC present sequence is a bisulphite treated human gene associated with
CC metastasis. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from USPTO at seqdata.uspto.gov/sequence.html?docID=20030148327.
XX
XX Sequence 14551 BP; 4362 A; 107 C; 3194 G; 6888 T; 0 U; 0 Other;
SQ
Query Match 14.7%; Score 43.2; DB 7; Length 14551;
Best Local Similarity 52.7%; Pred. No. 4.1;
Matches 117; Conservative 0; Mismatches 103; Indels 2; Gaps 1;
OY 57 TACTTCAAATTTATCTTCCAGTGGTTTGTATCTTCTCAATGTTATAGTGAATATAAT 116
Db 1078 TAATACTAAACACTATATAAAAAATTTTCTCTATTTAATAACAATCTAAATAAAAA 1019
OY 117 AATACATGCAGAAAAAGCATAAATCAACTATTGATTA--CGAAAAATGTGAACCTCTAT 174
Db 1018 CATACATTAATAATAAAAAATCAAAAAATTTAATAAAAAATTTCTAAATTTCTAAATAAAT 959
```

Qy	175	GTAAGTCAACTCAGGCTGCCTACCGTTTAAATTTCTCTCTAATCTTTAAACAAATA	234
Db	958	TCTATATAAATCCTCTCTAGCCCAATTTCTTTCACCTAAACGTCATAATATAAC	899
Qy	235	CTATCTTTTACTAAAGATGTTTCAGAGTAAGATAAATACAAA	276
Db	898	TTATTTAAATCAAAATCTTTCAAATATAATTAATAAATAAAA	857

Search completed: December 7, 2005, 00:02:57  
Job time : 465.191 secs





Qy	211	CTCTCTAAATCTTTAAAAA	CAAAATAC	TATCTTT	TACTTA	AGATGTTT	CAGAGTA	AGATAAAA	270
Db	11892	ATAATAATTTATTTAAAT	TCAACT	TACTTAAT	TACAAAAA	ACCTAAAA	ACATAAA	TAAATA	11833
Qy	271	TACAAAGAGAGAGTAAT	TCCCAAAA	294					
Db	11832	TACAATATCAATTAAT	TCAAAAAA	11809					

RESULT 5  
US-10-240-452-21/c  
; Sequence 21, Application US/10240452  
; Publication NO. US20030162194A1  
; GENERAL INFORMATION:  
; APPLICANT: OLEK, Alexander  
; APPLICANT: PIESENBRÖCK, Christian  
; APPLICANT: BERLIN, Kurt  
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Apoptosis  
; FILE REFERENCE: 5013.1006  
; CURRENT APPLICATION NUMBER: US/10/240.452  
; CURRENT FILING DATE: 2002-10-02  
; PRIOR APPLICATION NUMBER: PCT/EP01/03969  
; PRIOR FILING DATE: 2001-04-06  
; PRIOR APPLICATION NUMBER: DE 10019058.8  
; PRIOR FILING DATE: 2000-04-06  
; PRIOR APPLICATION NUMBER: DE 10019173.8  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: DE 10032529.7  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: DE 10043826.1  
; PRIOR FILING DATE: 2000-09-01  
; NUMBER OF SEQ ID NOS: 78  
; SEQ ID NO 21  
; LENGTH: 14924  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (208..209, 211, 213, 227)  
US-10-240-452-21

RESULT 6  
US-10-311-455-992/c  
; Sequence 992, Application US/10311455  
; Publication No. US20030143606A1

```

; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Imm
; TITLE OF INVENTION: cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; PRIOR FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 952
; LENGTH: 5311
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; US-10-311-455-992
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

```

```

RESULT 7
US-10-029-386-12125
; Sequence 12125, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC AC
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 12125
; LENGTH: 510
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHRL3
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.54
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.6
; OTHER INFORMATION: EST_HUMAN HIT: BF0632234.1, EVALUAE 1.00e-118
; OTHER INFORMATION: SWISSPROT HIT: P12270, EVALUAE 1.00e-33
; OTHER INFORMATION: NT HIT: g14507658, EVALUAE 1.00e-118
US-10-029-386-12125

```

Query Match	14.8%;	Score 43.6;	DB 6;	Length 510;
Best Local Similarity	48.8%;	Pred. No. 3.1;		
Matches 118;	Conservative 0;	Mismatches 124;	Indels 0;	Gaps 0;
Qy	2	GGCCTTAAAGGTTCTCTGCACAGTTGTCTCAAGAAGATCAGGATTTGAGCATCTCTTACTTT	61	
Db	225	GGCATCTAAAGTCCCTCTCTGGATGGTTTCCTTTTCAGTTCTTACTTTATTAGATCTTCCTT	284	
Qy	62	CAAAATTATCTTCCAGTGGTTTTGTTATCTTCTCAATGTTTATAGTGAATAATAATAATAC	121	
Db	285	CAGACTCTGAATTAAAGTCTGGTTGTAGTCAAAAGATGCATTTGATCTTTAAAAAATGGA	344	
Qy	122	ATGCAGAAAAAAGCATAAATCAACTATTTGAATAACGAAAAATGTCAAACCTCTATGCTAACTG	181	
Db	345	GGAAGAAGAAAGATTAAATATATAGACACATTTTTTCAAAGTTTGAAGTCCCAATGTAACCTT	404	
Qy	182	CAACTCAGGTCGCCACGGTTTTTTAAATTTCTCTCTAATCTTTAAAAACAATACTATCTTT	241	
Db	405	TATGTTATACCTGCTTTAATAACTTTTTTTTTTGAATAATCCTAAAAAATACCCATATACTTT	464	
Qy	242	TA	243	
Db	465	CA	466	

```

RESULT 8
US-09-925-065A-467258
; Sequence 467258, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 467258
; LENGTH: 646
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-467258

```

Query Match	14.8%	Score 43.6;	DB 4;	Length 646;
Best Local Similarity	47.4%;	Pred. No. 3.4;		
Matches 127;	Conservative 1;	Mismatches 140;	Indels 0;	Gaps 0;

  

Qy	20	ACAGTTCCTCAAGAAGATCAGGATTTGAGCATCTCTTACTTCAAATTTATTTCTTCCAGTG	79
Db	251	AGACTAATGGCCAAAGTACATTTTATTAAATCGTCTCTACTAATTGATTTTTTAAATGA	310
Qy	80	GTITTTGTATCTTCTCAATGTTATAGTGAATATATAATAATACATGCGAGAAAAGCATAA	139
Db	311	ATAAATCAATTTAAAAAATGCGAAAGATGATTTGTATTAATTCACACAGAAGAAATACAA	370
Qy	140	ATCAACTATTGGAATAACGAAATGTGAACCTCTATGTAACTTGCAACTCAGGCTCGCTTACC	199
Db	371	AYGGAAAACCAAGACACTTGGTCCATTATGTGAATATGAAAAATGCAAGTAAAAAACCAACAATA	430
Qy	200	GTITTTTTAATTTCTCTAAATCTTTTAAAAACAAATACTATCTTTTACTTAAAGATGTTTTCAGA	259

D<sub>b</sub> 431 AAATATAACTTTTAAACAAAATTTTCAAAACTTAAGTATTGGTAATATATATATTGTTGGTG 490  
Q<sub>y</sub> 260 GTAAGTAAATAACAAGAGAGTAATTC 287  
D<sub>b</sub> 491 GGAATTTTAAATTTTTAAAAATAGGAATC 518

```

RESULT 9
US-10-311-455-874/c
; Sequence 874, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311.455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 874
; LENGTH: 5204
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-874

```

Query Match	14.8%;	Score 43.4;	DB 6;	Length 5204;
Best Local Similarity	50.2%;	Pred. No. 8.5;		
Matches 107;	Conservative 0;	Mismatches 106;	Indels 0;	Gaps 0;
Qy	82	TTTGGTTATCTCTCAATGTGTATAGTGAATATATATAATACATGACAGAAAAAGCATAAAT	141	
Db	2290	TTTTTTTACAGGTAAACGCCCTTAATAACCAAAAAAAAAAAAAAAAAAAAAA	2231	
Qy	142	CAACTATTGAATTACGAAATGTGAACCTCTATGTAACTGCAACTCAGGCGCTACCGT	201	
Db	2230	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACAAAAAATTAATAATCATCAAAACCA	2171	
Qy	202	TTTTTAATTTCTCTCTAATCTTTAAAAACAATACTATCTTTTACTAAAGATGTTTCAGAGT	261	
Db	2170	TAAATTCATTCTCTTAAATTCCTATACTACATATAATTTTAAAAACGGTTCACAACT	2111	
Qy	262	ARGATAAAATACAAAGAGAGTAATTCCTCAAAA	294	
Db	2110	AAAAAAAAAAATAACTATAATAATAAACCCCAATA	2078	

```

RESULT 10
US-10-240-453-132/c
; Sequence 132, Application US/10240453
; Publication No. US20030148326A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA
; TITLE OF INVENTION: Transcription
; TITLE OF INVENTION: by Means of Assessing the Methylation Status of Genes
; TITLE OF INVENTION: With DNA Transcription
; FILE REFERENCE: 5013.1009
; CURRENT APPLICATION NUMBER: US/10/240,453
; CURRENT FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: PCT/EP01/03973
; PRIOR FILING DATE: 2001-04-06

```



```

; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 350
; SEQ ID NO 132
; LENGTH: 8170
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-240-453-132

Query Match      14.8%; Score 43.4; DB 6; Length 8170;
Best Local Similarity 50.2%; Pred. No. 10;
Matches 107; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

Qy      82 TTTGTTATCTCTCAATGTTTATAGTGAATATATAATATACATGCAGAAAAAGCATAAAT 141
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2628 TTTTAAATCTAAAACCTCTTTTAAATAAAATATTTTATAAAATATAAAACAAATAAAATAA 2569
Qy      142 CAACCTATTGAATAACGAAATGTGAACCTCTATGTAACCTGCAACTCAGGCTGCCTACCGT 201
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2568 ATATATATAAATAACAATATTTTAAAAATACTACTCTTTTATCTTAAAAATTTATTCGAT 2509
Qy      202 TTTTAAATTTCTCTTAATCTTTTAAAAACAATACTATCTTTTACTAAAGATGTTTCAGAGT 261
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2508 TATTAAATTTACATTTACTATCTTAAAAACAACCTATAAACCCTTATATTAACACAACAAAC 2449
Qy      262 AAGATAAATAACAAGAGAGTAATTTCCCAAAA 294
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2448 AATCTAAATATACACAACAATAAATTTTCATAAA 2416

RESULT 11
US-10-311-455-1380/c
; Sequence 1380, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation
; FILE REFERENCE: 5013.1014
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: US/10/311,455
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1380
; LENGTH: 8423
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 6659
; OTHER INFORMATION: n is a or g or c or t
US-10-311-455-1380

Query Match      14.8%; Score 43.4; DB 6; Length 8423;
Best Local Similarity 51.9%; Pred. No. 10;
Matches 98; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

```

```

Qy      86 TTATCTTCTCAATGTTTATAGTGAATATATAATATACATGCAGAAAAAGCATAAATCAAC 145
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1174 TCATCCACAAAAACATAACCTCCCAATATATACTCTCAATATATAAAACAACTAAAAATAAC 1115
Qy      146 TATTGAATAACGAAATGTGAACCTCTATGTAACCTGCAACTCAGGCTGCCTACGGTTTTT 205
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1114 AAAAAACATTTTAAAAATAAAACCTTTTAAATTTAAATTTAAATTTAAATTTAAATTTAAAT 1055
Qy      206 AATTCTCTCTTAATCTTTTAAAAACAATACTATCTTTTACTAAAGATGTTTCAGAGTAAGA 265
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1054 ACTTCTCCAAATATTTTAAAAACACAAATACCATATAAAAAAATAAATAAATTTTATA 995
Qy      266 TAAAAATACA 274
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
994 TAAATTACA 986

RESULT 12
US-10-312-841-1/c
; Sequence 1, Application US/10312841
; Publication No. US20030186277A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MHC
; FILE REFERENCE: E01/1208/WO
; CURRENT APPLICATION NUMBER: US/10/312,841
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 2
; SEQ ID NO 1
; LENGTH: 3673778
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; NAME/KEY: unsure
; LOCATION: (3294164)
US-10-312-841-1

Query Match      14.8%; Score 43.4; DB 6; Length 3673778;
Best Local Similarity 50.2%; Pred. No. 1e+02;
Matches 107; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

Qy      82 TTTGTTATCTCTCAATGTTTATAGTGAATATATAATATACATGCAGAAAAAGCATAAAT 141
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1560380 TTTTITTTACAGTAAACGCCCTATAAACHAAAAAANAANAANAANAANAANAANAANA 1560321
Qy      142 CAACCTATTGAATAACGAAATGTGAACCTCTATGTAACCTGCAACTCAGGCTGCCTACCGT 201
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1560320 AAAAAAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANA 1560261
Qy      202 TTTTAAATTTCTCTTAATCTTTTAAAAACAATACTATCTTTTACTAAAGATGTTTCAGAGT 261
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1560260 TAAATTCATTTTCTTAAAAATCCCTATACATATAATTTTAAAAAANAACGTTTCTACAACT 1560201
Qy      262 AAGATAAATAACAAGAGAGTAATTTCCCAAAA 294
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1560200 AAAAAAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANA 1560168

RESULT 13
US-10-240-485-137/c
; Sequence 137, Application US/10240485
; Publication No. US20030148327A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with
; FILE REFERENCE: 5013.1007
; CURRENT APPLICATION NUMBER: US/10/240,485

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 11:34:11 ; Search time 200.691 Seconds  
(without alignments)  
455.919 Million cell updates/sec

Title: US-09-980-046B-2

Perfect score: 294

Sequence: 1 cggccttaaggcttcctctga.....aagagagtaattcccaaaaa 294

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3289935 seqs, 155610033 residues

Total number of hits satisfying chosen parameters: 6579870

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications\_NA\_New.\*

1: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	39.4	13.4	4497	US-10-518-753-15	Sequence 15, Appl
C 2	36.8	12.5	2467	US-10-750-185-29248	Sequence 29248, A
C 3	36	12.2	1857	US-10-750-185-42217	Sequence 42217, A
C 4	35.6	12.1	1740	US-10-750-185-37198	Sequence 37198, A
C 5	35.6	12.1	2536	US-11-087-227-9	Sequence 9, Appli
C 6	35.6	12.1	2613	US-11-087-227-7	Sequence 7, Appli
C 7	35.6	12.1	2748	US-11-087-227-5	Sequence 5, Appli
C 8	35.4	12.0	1899	US-10-750-185-62020	Sequence 62020, A
C 9	35.2	12.0	2444	US-10-750-185-29713	Sequence 29713, A
C 10	35.2	12.0	5501	US-10-750-185-28699	Sequence 28699, A
C 11	35	11.9	3069	US-10-750-185-4352	Sequence 4352, Ap
C 12	35	11.9	3206	US-10-750-185-4382	Sequence 4382, Ap
C 13	34.8	11.8	1352	US-11-096-622-17	Sequence 17, Appl
C 14	34.8	11.8	1757	US-11-096-622-14	Sequence 14, Appl
C 15	34.8	11.8	8512	US-11-096-622-20	Sequence 20, Appl
C 16	34.8	11.8	8565	US-11-096-622-21	Sequence 21, Appl
C 17	34.6	11.8	4731	US-10-750-185-60556	Sequence 60556, A
C 18	34.2	11.6	2253	US-10-750-185-6095	Sequence 6095, Ap
C 19	34.2	11.6	2311	US-10-485-517-68	Sequence 485, Appl
C 20	34.2	11.6	3416	US-10-750-185-3946	Sequence 3946, Ap
C 21	34.2	11.6	3520	US-10-750-185-3657	Sequence 3657, Ap
C 22	34	11.6	906	US-10-750-185-205	Sequence 205, App
C 23	34	11.6	2625	US-10-750-185-37777	Sequence 37777, A

#### ALIGNMENTS

##### RESULT 1

US-10-518-753-15/c

; Sequence 15, Application US/10518753

; Publication No. US20050262589A1

; GENERAL INFORMATION:

; APPLICANT: Calgene LLC

; TITLE OF INVENTION: Nucleic Acid Sequences and Methods of Use for the Production of P

; FILE REFERENCE: 16518.129

; CURRENT APPLICATION NUMBER: US/10/518.753

; CURRENT FILING DATE: 2004-12-21

; NUMBER OF SEQ ID NOS: 39

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 15

; LENGTH: 4497

; TYPE: DNA

; ORGANISM: Glycine max

US-10-518-753-15

Query Match 13.4%; Score 39.4; DB 6; Length 4497;

Best Local Similarity 55.4%; Pred. No. 1.5;

Matches 97; Conservative 0; Mismatches 76; Indels 2; Gaps 1;

Qy	106	TGAATAATATATACATGCAGAAAGCAATCACTATTGAATAACGA--AAATG	163
Db	3336	TGATATATATCAAGACATATATATAAAATTAAGTTTAAATTAGATTAATTAAT	3277
Qy	164	TGAACCTCTATGTAACTGCAACTCAGGCTGCCTACCGTTTTTAATTTCTCTAATCTTT	223
Db	3276	TGAATCTAAATACATGTCCTAATTTAGTGTGCTATATTTTATTTTCAATGACATT	3217
Qy	224	AAAAACAAATCTACTATCTTTTACTAAAGATGTTTCAGAGTAAGATAAATACAAAGA	278
Db	3216	TATTACAAATTTTTCAGTTTAAATGAATCAATTTGATTTTAAATACCTTGTAATA	3162

##### RESULT 2

US-10-750-185-29248

; Sequence 29248, Application US/10750185

; Publication No. US20050260603A1

; GENERAL INFORMATION:

; APPLICANT: MMI GENOMICS, INC.

; APPLICANT: DENISE, Sue K.

; APPLICANT: KERR, Richard

; APPLICANT: ROSENFELD, David

; APPLICANT: HOLM, Tom



```
; LENGTH: 2536
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-087-227-9

Query Match      12.1%; Score 35.6; DB 7; Length 2536;
Best Local Similarity 55.7%; Pred. No. 9.7;
Matches 68; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

Qy 159 AAATGTGAACCTCTATGTAACTGCAACTCAGGCTGCCTACCGTTTAAATTTCTCTCTAA 218
    |||||
Db 2415 AATTGTGACATATGTATTATATAAACTTAACCTTTTAACTACTGTTTATTTTAGCCCAT 2474
    |||||

Qy 219 TCTTTAAAAACAATACTATCTTTTAAAGATGTTTCAGAGTAAGATAAAATACAAGA 278
    |||||
Db 2475 TGTTTAAAAATAAAGTTAAAAAAATTTAACTGCTTAAAGTAAGAAAAAATAAAAAA 2534
    |||||

Qy 279 GA 280
    |
Db 2535 AA 2536

RESULT 6
US-11-087-227-7
; Sequence 7, Application US/11087227
; Publication No. US20050260566A1
; GENERAL INFORMATION:
; APPLICANT: Fischer, Timothy J.
; APPLICANT: Malinowski, Douglas P.
; APPLICANT: Taylor, Adriann J.
; APPLICANT: Parker, Margaret R.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE
; TITLE OF INVENTION: DETECTION OF CERVICAL DISEASE
; FILE REFERENCE: 046143/287139
; CURRENT APPLICATION NUMBER: US/11/087,227
; PRIOR FILING DATE: 2005-03-23
; PRIOR FILING DATE: 2005-03-23
; PRIOR FILING DATE: 2004-03-24
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 2613
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-087-227-7

Query Match      12.1%; Score 35.6; DB 7; Length 2613;
Best Local Similarity 55.7%; Pred. No. 9.8;
Matches 68; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

Qy 159 AAATGTGAACCTCTATGTAACTGCAACTCAGGCTGCCTACCGTTTAAATTTCTCTCTAA 218
    |||||
Db 2492 AATTGTGACATATGTATTATATAAACTTAACCTTTTAACTACTGTTTATTTTAGCCCAT 2551
    |||||

Qy 219 TCTTTAAAAACAATACTATCTTTTAAAGATGTTTCAGAGTAAGATAAAATACAAGA 278
    |||||
Db 2552 TGTTTAAAAATAAAGTTAAAAAAATTTAACTGCTTAAAGTAAGAAAAAATAAAAAA 2611
    |||||

Qy 279 GA 280
    |
Db 2612 AA 2613

RESULT 7
US-11-087-227-5
; Sequence 5, Application US/11087227
; Publication No. US20050260566A1
; GENERAL INFORMATION:
; APPLICANT: Fischer, Timothy J.
; APPLICANT: Malinowski, Douglas P.
; APPLICANT: Taylor, Adriann J.
; APPLICANT: Parker, Margaret R.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE
```

```
; TITLE OF INVENTION: DETECTION OF CERVICAL DISEASE
; FILE REFERENCE: 046143/287139
; CURRENT APPLICATION NUMBER: US/11/087,227
; CURRENT FILING DATE: 2005-03-23
; PRIOR FILING DATE: 2005-03-23
; PRIOR FILING DATE: 2004-03-24
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 2748
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-087-227-5

Query Match      12.1%; Score 35.6; DB 7; Length 2748;
Best Local Similarity 55.7%; Pred. No. 9.9;
Matches 68; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

Qy 159 AAATGTGAACCTCTATGTAACTGCAACTCAGGCTGCCTACCGTTTAAATTTCTCTCTAA 218
    |||||
Db 2627 AATTGTGACATATGTATTATATAAACTTAACCTTTTAACTACTGTTTATTTTAGCCCAT 2686
    |||||

Qy 219 TCTTTAAAAACAATACTATCTTTTAAAGATGTTTCAGAGTAAGATAAAATACAAGA 278
    |||||
Db 2687 TGTTTAAAAATAAAGTTAAAAAAATTTAACTGCTTAAAGTAAGAAAAAATAAAAAA 2746
    |||||

Qy 279 GA 280
    |
Db 2747 AA 2748

RESULT 8
US-10-750-185-62020
; Sequence 62020, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: WMI1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 62020
; LENGTH: 1899
; TYPE: DNA
; ORGANISM: Bovine
US-10-750-185-62020

Query Match      12.0%; Score 35.4; DB 6; Length 1899;
Best Local Similarity 54.1%; Pred. No. 10;
Matches 72; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

Qy 113 TAATAATACATGCAGAAAAAGCATAAATCAACTATTGAAATACGAAATGTGAACCTCT 172
    |||||
Db 1079 TAATGATAAGGGAAAAAAATTAATAAAAAAGATGATAAAGGAAGCTTCAAAACCT 1138
    |||||

Qy 173 ATGTAACCTGCAACTCAGGCTGCCTACCGTTTAAATTTCTCTCTAAATTTTAAACAAA 232
    |||||
Db 1139 ATGTACATCTCTCTCAAGAAATCCCTTCTAGATTTTGTGTATCTGTCTACTTTCACACACTA 1198
    |||||

Qy 233 TACTATCTTTACT 245
    |||||
Db 1199 AACATATAACCT 1211
```

```
RESULT 9
US-10-750-185-29713/c
; Sequence 29713, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR FILING DATE: 2003-12-31
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 29713
; LENGTH: 2444
; TYPE: DNA
; ORGANISM: Bovine 19866880767225
US-10-750-185-29713

Query Match      12.0%; Score 35.2; DB 6; Length 2444;
Best Local Similarity 48.5%; Pred. No. 12;
Matches 97; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

Qy 34 AAGATCAGGATTTGAGCATCTTACTTCAAAATTATCTTCCAGTGGTTTGTATCTTC 93
Db 2195 AAGCAAAGATCCAAGCTCTCTCTTAACTTTTAGATTGATGGTTCTCTGTGCC 2136

Qy 94 TCAATGTTATAGTGAATATAAATACATGCGAGAAAGAGCAATAATCACTATTGAAT 153
Db 2135 ACAATATAAATCATGCTTTCAGGATCTTCATCTTGGGAAAGAAAGAGAGAAACAT 2076

Qy 154 AAGCAAAATGTGAACCTCTATGTAACCTGCAACTCAGCTGCCCTACCGTTTTTAATTCTC 213
Db 2075 CCATCCAGGTCTCCCTTTATCTGTTGCCCTTCTTCGCACATGCACCTTGTCTCTA 2016

Qy 214 TCTAATCTTTAAACAAAT 233
Db 2015 TTTCTTCTCAAAAGAAAAT 1996

RESULT 10
US-10-750-185-28699
; Sequence 28699, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR FILING DATE: 2003-12-31
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 28699
; LENGTH: 5501
; TYPE: DNA
; ORGANISM: Bovine 19866880637762
US-10-750-185-28699

Query Match      12.0%; Score 35.2; DB 6; Length 5501;
Best Local Similarity 52.8%; Pred. No. 15;
Matches 76; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

Qy 137 TAAATCAACTATTGAAATAACGAAATGTGAACCTCTATGTAACCTCAACTCAGGCTGCCT 196
Db 4128 TAGATAAACTAATGAACAACACAGACAAAGACTCTTATGGAACCTTAATGCACAGAAAT 4187

Qy 197 ACCGTTTTTAATTTCTCTCTAATCTTTAAAAACAAATATCTATCTTTACTAAAGATGTTTC 256
Db 4188 GTATATTTTAAAAATCCTAAAACTATACATTAATTAGCAATACATATATTATCATTTCTC 4247

Qy 257 AGAGTAAAGATAAAATACAAAGAGA 280
Db 4248 AATACAGAAAAAAGACATAATGA 4271

RESULT 11
US-10-793-626-4352
; Sequence 4352, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4352
; LENGTH: 3069
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-4352

Query Match      11.9%; Score 35; DB 6; Length 3069;
Best Local Similarity 55.3%; Pred. No. 14;
Matches 68; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

Qy 65 ATTATTCTTCCAGTGGTTTGTGTTTATCTTCTCAATGTTATAGTGAATATAATATACATG 124
Db 2418 ACTGTCCACCTGTGCTTTTATTAAAAATTTATTAGTAAATAAAAAATATAAAATTTACAG 2477

Qy 125 CAGAAAAAGCATAAATCAACTATTGCAATACGAAAAATGGAACCTCTATGTAACCTCAA 184
Db 2478 AAAAAAATATCAAAAAACCAAAATATAAAAAATGAAATGATGGATTAGGATGCTTTTGCTA 2537

Qy 185 CTC 187
Db 2538 CAC 2540

RESULT 12
US-10-793-626-4382/c
; Sequence 4382, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4382
```

```
; LENGTH: 3206
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-09-980-046b-2
Query Match      11.9%; Score 35; DB 6; Length 3206;
Best Local Similarity 55.3%; Pred. No. 14;
Matches 68; Conservative 0; Mismatches 55; Indels 0; Gaps 0;
Qy 65 ATTATCTTCAGCGTGGTTTCTTCAATGTTATAGTGAATAATAATAATCATG 124
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1064 ACTGCCACCTTGCTTTTATTAATAAATATTATTAGTAAATAATAATAATACAG 1005
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 125 CAGAAAACGATTAATCACTATTGTATACGAAATGTGAACCTCTATGTAACGCA 184
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1004 AAAAAAATATCAAAACCAAAATAATAAAATGAATGATGATGATGATGATGATG 945
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 185 CTC 187
    |||
Db 944 CAC 942
    |||
RESULT 13
US-11-096-622-17/c
; Sequence 17, Application US/11096622
; Publication No. US20050260643A1
; GENERAL INFORMATION:
; APPLICANT: HUNG, MIEN-CHIE
; APPLICANT: DAY, CHI-PING
; APPLICANT: RAU, KUN-MING
; APPLICANT: XIE, XIAOMING
; APPLICANT: LI, ZHENG
; TITLE OF INVENTION: CANCER SPECIFIC PROMOTERS
; FILE REFERENCE: UTSC:845US
; CURRENT APPLICATION NUMBER: US/11/096,622
; CURRENT FILING DATE: 2005-04-01
; PRIOR APPLICATION NUMBER: 60/559,111
; PRIOR FILING DATE: 2004-04-02
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 1352
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Primer
US-11-096-622-17
Query Match      11.8%; Score 34.8; DB 7; Length 1352;
Best Local Similarity 50.6%; Pred. No. 13;
Matches 84; Conservative 0; Mismatches 82; Indels 0; Gaps 0;
Qy 78 TGGTTTCTTATCTTCTCAATGTTATAGTGAATAATAATAATAATAATAATGCA 137
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1267 TTGTTTCGTTTGTGTTGATGAGCGGTATGTTAGTACTATCGATTACACAAAAAACC 1208
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 138 AAATCAACTATTGAATAACGAAATGTGAACCTCTATGTAACCTCAGGCTGCCTA 197
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1207 AACACACAGATGTAATGAATAAAGATATTTTATGCGCGCGCTCTAGAAATCTA 1148
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 198 CGTTTTTAATTTCTCTCTAATCTTTTAAACAAATACTATCTTTA 243
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1147 GCTGATATAGTGTGGCTCAAAACCTTCAGCACAATCACACCGTTA 1102
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
RESULT 14
US-11-096-622-14/c
; Sequence 14, Application US/11096622
; Publication No. US20050260643A1
```

```
; GENERAL INFORMATION:
; APPLICANT: HUNG, MIEN-CHIE
; APPLICANT: DAY, CHI-PING
; APPLICANT: RAU, KUN-MING
; APPLICANT: XIE, XIAOMING
; APPLICANT: LI, ZHENG
; TITLE OF INVENTION: CANCER SPECIFIC PROMOTERS
; FILE REFERENCE: UTSC:845US
; CURRENT APPLICATION NUMBER: US/11/096,622
; CURRENT FILING DATE: 2005-04-01
; PRIOR APPLICATION NUMBER: 60/559,111
; PRIOR FILING DATE: 2004-04-02
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 1757
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Primer
US-11-096-622-14
Query Match      11.8%; Score 34.8; DB 7; Length 1757;
Best Local Similarity 50.6%; Pred. No. 14;
Matches 84; Conservative 0; Mismatches 82; Indels 0; Gaps 0;
Qy 78 TGGTTTCTTATCTTCTCAATGTTATAGTGAATAATAATAATAATAATAATGCA 137
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1672 TTGTTTCGTTTGTGTTGATGAGCGGTATGTTAGTACTATCGATTACACAAAAAACC 1613
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 138 AAATCAACTATTGAATAACGAAATGTGAACCTCTATGTAACCTCAGGCTGCCTA 197
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1612 AACACACAGATGTAATGAATAAAGATATTTTATGCGCGCGCTCTAGAAATCTA 1553
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 198 CGTTTTTAATTTCTCTCTAATCTTTTAAACAAATACTATCTTTA 243
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1552 GCTGATATAGTGTGGCTCAAAACCTTCAGCACAATCACACCGTTA 1507
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
RESULT 15
US-11-096-622-20/c
; Sequence 20, Application US/11096622
; Publication No. US20050260643A1
; GENERAL INFORMATION:
; APPLICANT: HUNG, MIEN-CHIE
; APPLICANT: DAY, CHI-PING
; APPLICANT: RAU, KUN-MING
; APPLICANT: XIE, XIAOMING
; APPLICANT: LI, ZHENG
; TITLE OF INVENTION: CANCER SPECIFIC PROMOTERS
; FILE REFERENCE: UTSC:845US
; CURRENT APPLICATION NUMBER: US/11/096,622
; CURRENT FILING DATE: 2005-04-01
; PRIOR APPLICATION NUMBER: 60/559,111
; PRIOR FILING DATE: 2004-04-02
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 8512
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Primer
US-11-096-622-20
Query Match      11.8%; Score 34.8; DB 7; Length 8512;
Best Local Similarity 50.6%; Pred. No. 20;
Matches 84; Conservative 0; Mismatches 82; Indels 0; Gaps 0;
Qy 78 TGGTTTCTTATCTTCTCAATGTTATAGTGAATAATAATAATAATAATAATGCA 137
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Db 2952 TTGTTTCGTTTTGTTTTGATGGAGCGGTATGTTAGTACTATCGATTTCACACAAAAACC 2893  
Qy 138 AAATCAACTATTGAATAACGAAATGTGAACCTCTATGTAACTGCAACTCAGGCTGCCTA 197  
Db 2892 AACACACAGATGTAATGAAAAATAAGATAATTTTATTGCGGCCGCCCTCTAGAAATCTA 2833  
Qy 198 CCGTTTTTAATTTCTCTCTAATCTTTAAAAACAAATACTATCTTTA 243  
Db 2832 GCTGATATAGTGGCTCAAAACCTTCAGCACAANTCACCCTTA 2787

Search completed: December 6, 2005, 23:16:50  
Job time : 201.691 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 11:47:47 ; Search time 124.194 Seconds  
(without alignments)  
4239.411 Million cell updates/sec

Title: US-09-980-046B-3

Perfect score: 79

Sequence: 1 cgggaatcccccttcagttc.....aaacaacgactctaaaaa 79

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N\_Geneseq\_21.\*

- 1: Geneseqn1980s.\*
- 2: Geneseqn1990s.\*
- 3: Geneseqn2000s.\*
- 4: Geneseqn2001as.\*
- 5: Geneseqn2001bs.\*
- 6: Geneseqn2002as.\*
- 7: Geneseqn2002bs.\*
- 8: Geneseqn2003as.\*
- 9: Geneseqn2003bs.\*
- 10: Geneseqn2003cs.\*
- 11: Geneseqn2003ds.\*
- 12: Geneseqn2004as.\*
- 13: Geneseqn2004bs.\*
- 14: Geneseqn2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	79	100.0	79	4	AAC89734
2	75.8	95.9	1464	3	AAC59772 Human sec
3	74.8	94.7	428	13	ADU14748
4	74.8	94.7	522	14	Adu14748 Solid tum
5	74.8	94.7	592	10	Adf81837 Leukaemia
6	74.8	94.7	1985	12	Adq86642 Human tum
7	74.8	94.7	1985	13	Adq85525 Human tum
8	74.8	94.7	2000	4	AAC88152
9	74.8	94.7	2000	6	ABL65927 Lung canc
10	74.8	94.7	2000	6	ABK84431 Human CDN
11	74.8	94.7	2000	12	Adj74845 Marker ge
12	74.8	94.7	2000	13	Adr52746 Drug ther
13	74.8	94.7	2000	13	Adp24663 PRO polyp
14	74.8	94.7	2000	13	Adu06261 Novel bro
15	74.8	94.7	2000	14	Adw38555 Immunomod
16	74.8	94.7	2180	2	AAV32446 K12 CDNA
17	71	89.9	90	4	AAC89735 Human gas
18	63.4	80.3	212	2	AAT26356 Human gen
19	54	68.4	401	10	ADF80286 Leukaemia

20	29.6	37.5	110000	6	ABA03041_08	Continuation (9 of
21	29.2	37.0	2278	6	ABQ70619	Abq70619 Listeria
22	29	36.7	528	5	ADL43615	Adl43615 Human ova
23	28.8	36.5	16037	4	AAK78529	AAk78529 Human imm
24	28.8	36.5	16037	6	ABK69947	ABk69947 Human sec
25	28.8	36.5	16037	8	ABZ74623	ABz74623 Secreted
26	28.8	36.5	16037	10	ADC21014	Adc21014 Human sec
27	28.8	36.5	16037	10	ABZ68144	ABz68144 Human sec
28	28.8	36.5	16107	4	AAK78528	AAk78528 Human imm
29	28.8	36.5	16107	6	ABK69948	ABk69948 Human sec
30	28.8	36.5	16107	8	ABZ74624	ABz74624 Secreted
31	28.8	36.5	16107	10	ADC21015	Adc21015 Human sec
32	28.8	36.5	16107	10	ABZ68145	ABz68145 Human sec
33	28.8	36.5	31828	11	ACN44338	ACn44338 Human gen
34	28.4	35.9	110000	14	ADZ13631_2	Continuation (3 of
35	28.4	35.9	110000	14	ADZ13631_3	Continuation (4 of
36	28.4	35.9	110000	14	ADZ13620_2	Continuation (3 of
37	28.4	35.9	110000	14	ADZ13620_3	Continuation (4 of
38	28.4	35.9	299598	12	ADQ59380	Adq59380 Human can
39	28	35.4	683	5	ADL37231	Adl37231 Human ova
40	28	35.4	683	5	ADI72082	Adi72082 Human ova
41	28	35.4	4722	14	ADW10342	Adw10342 Colon pro
42	28	35.4	17144	4	AAS46665	Aas46665 Tumour su
43	27.8	35.2	495	4	ABA55327	Aba55327 Human foe
44	27.8	35.2	495	4	AAI34978	AAi34978 Probe #36
45	27.8	35.2	495	4	ABA44865	Aba44865 Human bre

#### ALIGNMENTS

RESULT 1

AAC89734

ID AAC89734 standard; CDNA; 79 BP.

XX AAC89734;

XX 12-MAR-2001 (first entry)

XX Human gastrointestinal inflammation-related CDNA, SEQ ID NO: 3.

XX Human; cytostatic; immunomodulator; immunostimulant; vulnery;

XX anti-inflammatory; neuroprotective; antibacterial; gene therapy;

XX gastrointestinal inflammation; immune system disorder; genetic disorder;

XX cancer; autoimmune disorder; infection; wound healing; ss.

XX Homo sapiens.

XX WO200073324-A2.

XX 07-DEC-2000.

XX 01-JUN-2000; 2000WO-US015191.

XX 01-JUN-1999; 99US-0137058P.

XX (DIGI-) DIGITAL GENE TECHNOLOGIES INC.

XX Youakim A, Dubose RF, Sims JE, Pribyl TM, Hillbush BS, Hasel KW;

XX WPI; 2001-061508/07.

XX New polynucleotides and polypeptides, useful in gene therapy and in  
diagnosing a pathological condition, e.g. for modulating gene expression  
in gastrointestinal inflammation, or for treating cancers or genetic  
disorders.

XX Claim 1; Page 84; 108pp; English.

XX The present sequence is one of a number of isolated human polynucleotides  
which are useful in gene therapy, and for diagnosing a pathological  
condition or a susceptibility to it. In particular, the polynucleotides  
are useful for modulating gene expression in gastrointestinal

CC inflammation. The polynucleotides are useful for chromosome  
 CC identification, controlling gene expression through triple helix  
 CC formation or antisense DNA or RNA, or identifying individuals from minute  
 CC biological samples using DNA-based identification techniques. The  
 CC polynucleotides can also be used as an alternative to restriction  
 CC fragment length polymorphism (RFLP), by determining the actual base-by-  
 CC base DNA sequences of selected portions of an individual's genome. The  
 CC polynucleotides may also be used as molecular weight markers on Southern  
 CC gels, as diagnostic probes for the presence of a specific mRNA, as a  
 CC probe to subtract-out known sequences in the process of discovering novel  
 CC polynucleotides, or as an antigen to elicit an immune response. The  
 CC polypeptides are useful in diagnostic procedures to detect a disorder.  
 CC The polynucleotides and polypeptides are useful for preventing, treating  
 CC or ameliorating immune system disorders, genetic disorders, cancers, some  
 CC autoimmune disorders, or infections. The polynucleotides and polypeptides  
 CC are also useful for differentiating, proliferating or attracting cells,  
 CC leading to the regeneration of tissues, especially in wounds or burns.  
 CC The polypeptides and polynucleotides may also be used as a food additive  
 CC or preservative  
 XX  
 SQ Sequence 79 BP; 31 A; 18 C; 12 G; 18 T; 0 U; 0 Other;

Query Match 100.0%; Score 79; DB 4; Length 79;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-15;  
 Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60  
 Db 1 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60  
 Qy 61 AACAAACCGACTCTAAAAA 79  
 Db 61 AACAAACCGACTCTAAAAA 79

RESULT 2  
 AAC59772

ID AAC59772 standard; cDNA; 1464 BP.

AC AAC59772;

XX 26-JAN-2001 (first entry)

XX Human secreted protein gene 35 SEQ ID NO:45.

XX Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;  
 KW antirheumatic; antiproliferative; cytostatic; cardiant; vasotropic;  
 KW cerebroprotective; nontropic; neuroprotective; antibacterial; virucide;  
 KW fungicide; ophthalmological; gene therapy; autoimmune disease; neoplasm;  
 KW hyperproliferative disorder; cancer; cardiovascular disorder; infection;  
 KW cerebrovascular disorder; angiogenesis; nervous system disorder;  
 KW ocular disorder; wound healing; skin aging; food additive; preservative;  
 KW ss.

XX Homo sapiens.

XX W0200056751-A1.

XX 28-SEP-2000.

XX 09-MAR-2000; 2000WO-US006013.

XX 19-MAR-1999; 99US-0125360P.

XX 11-JUN-1999; 99US-0138626P.

XX 03-DEC-1999; 99US-0168662P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM, Komatsoulis G;

XX WPI; 2000-579482/54.

XX P-PSDB; AAB34611.

PT Isolated nucleic acid molecule encoding a human secreted protein is used  
 PT in preventing, treating or ameliorating a medical condition.  
 XX  
 PS Claim 1; Page 362; 419pp; English.

XX The polynucleotide sequences given in AAC59738 to AAC59787 encode the  
 CC human secreted proteins given in AAB34577 to AAB34626. AAB34627 to  
 CC AAB34686 represent human secreted polypeptide sequences and proteins  
 CC homologous to them, which are given in the exemplification of the present  
 CC invention. Human secreted proteins have activities based on the tissues  
 CC and cells the genes are expressed in. Example of activities include:  
 CC antiarthritic; immunosuppressive; antirheumatic; antiproliferative;  
 CC cytostatic; cardiant; vasotropic; cerebroprotective; nontropic;  
 CC neuroprotective; antibacterial; virucide; fungicide; and  
 CC ophthalmological. The polynucleotides and proteins can be used to  
 CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,  
 CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used  
 CC in diagnosing a pathological condition or susceptibility to a  
 CC pathological condition. Disorders which are diagnosed or treated include  
 CC autoimmune diseases, hyperproliferative disorders e.g. neoplasms and  
 CC cancers of the breast or liver, cardiovascular disorders, cerebrovascular  
 CC disorders, angiogenesis, nervous system disorders, infections caused by  
 CC bacteria, viruses and fungi and ocular disorders. The proteins can also  
 CC be used to aid wound healing and epithelial cell proliferation, to  
 CC prevent skin aging due to sunburn, to maintain organs before  
 CC transplantation, for supporting cell culture of primary tissues, to  
 CC regenerate tissues and in chemotaxis. The proteins can also be used as a  
 CC food additive or preservative to increase or decrease storage  
 CC capabilities. AAC59729 to AAC59737 and AAB34576 represent sequences used  
 CC in the exemplification of the present invention

XX Sequence 1464 BP; 309 A; 483 C; 408 G; 264 T; 0 U; 0 Other;

Query Match 95.9%; Score 75.8; DB 3; Length 1464;  
 Best Local Similarity 97.5%; Pred. No. 3.4e-14;  
 Matches 77; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60  
 Db 1071 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1130  
 Qy 61 AACAAACCGACTCTAAAAA 79  
 Db 1131 AACAAACCGACTCTAAAAA 1149

RESULT 3  
 ADU14748

ID ADU14748 standard; DNA; 428 BP.

XX ADU14748;

XX 27-JAN-2005 (first entry)

XX Solid tumour prognosis gene seqid 5187.

XX cytostatic; gene therapy; expression profile; solid tumour;  
 KW peripheral blood mononuclear cell; PBMC; prognosis; ds.

XX Unidentified.

XX W02004097052-A2.

XX 11-NOV-2004.

XX 29-APR-2004; 2004WO-US013587.

XX 29-APR-2003; 2003US-0466067P.

XX 23-JAN-2004; 2004US-0538246P.

XX (AMHP ) WYETH.

XX (STRA/) STRAHS A.

PI Strahs A, Trepicchio WL, Burczynski ME, Twine NC, Slonim DK;  
 PI Immerman F, Dorner AJ;  
 XX  
 DR WPI; 2004-804779/79.  
 XX  
 PT A method, useful for prognosing and treating solid tumor, comprises  
 PT comparing an expression profile of a gene expressed in peripheral blood  
 PT mononuclear cells to a reference expression profile of a gene.  
 XX  
 PS Disclosure; Page; 111pp; English.  
 XX  
 CC The invention describes a method comprising comparing an expression  
 CC profile of at least one gene in a peripheral blood sample of a patient to  
 CC at least one reference expression profile of the at least one gene, where  
 CC the patient has a solid tumor, and each of the gene is differentially  
 CC expressed in peripheral blood mononuclear cells (PBMCs) of a first class  
 CC of patients as compared to PBMCs of a second class of patients, where  
 CC both the first and second classes of patients have the solid tumor, and  
 CC each of the first and second classes is a subcluster formed by an  
 CC unsupervised clustering analysis of gene expression profiles in PBMCs of  
 CC a population of patients who have the solid tumor, and where the  
 CC majority of the first class of patients has a first clinical outcome, and  
 CC the majority of the second class of patients has a second clinical  
 CC outcome. Also described are: a system comprising (i) a memory or a  
 CC storage medium including data that represent an expression profile of at  
 CC least one gene in a peripheral blood sample of a patient who has a solid  
 CC tumor, (ii) at least another storage medium including data that  
 CC represent at least one reference expression profile of the gene, (iii) a  
 CC program capable of comparing the expression profile to the reference  
 CC expression profile, and (iv) a processor capable of executing the  
 CC program, where expression levels of the gene in peripheral blood  
 CC mononuclear cells of patients who have the solid tumor correlate with  
 CC clinical outcomes of the patients; and a nucleic acid or protein array  
 CC comprising concentrated probes for solid tumor prognosis genes, where  
 CC each of the solid tumor prognosis genes is differentially expressed in  
 CC PBMCs of a first class of patients as compared to PBMCs of a second class  
 CC of patients, where both the first and second classes of patients have a  
 CC solid tumor, and where the first class of patients has a first clinical  
 CC outcome, and the second class of patients has a second clinical outcome.  
 CC The method, system, and array are useful for prognosing and treating  
 CC solid tumors. This sequence represents a solid tumor prognosis gene of  
 CC the invention. Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic format  
 CC directly from WIPO at ftp.wipo.int/pub/intpublished\_pet\_sequences.  
 XX  
 SQ Sequence 428 BP; 107 A; 108 C; 112 G; 81 T; 0 U; 20 Other;  
 Query Match 94.7%; Score 74.8; DB 13; Length 428;  
 Best Local Similarity 97.4%; Pred. No. 5.5e-14;  
 Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 CGGGATCCCTTCAGTCTTTGAAAAGTTCATGACTCGAATATCTGAATGAAGAA 60  
 Db 100 CGGGATCCCTTCAGTCTTTGAAAAGTTCATGACTCGAATATCTGAATGAAGAA 159  
 QY 61 AACAAACCGACTCTAAA 78  
 Db 160 AACAAACCGACTCTACAAA 177  
 RESULT 4  
 ID AEA36210/c  
 XX ID AEA36210 standard; DNA; 522 BP.  
 AC AEA36210;  
 XX  
 DT 25-AUG-2005 (first entry)  
 XX  
 DE Human nucleic acid sequence #142.  
 XX  
 KW Screening; gene expression; colorectal tumor; colitis; Crohn's disease;  
 KW irritable bowel syndrome; gastrointestinal disease; cytostatic;  
 KW gastrointestinal-gen.; antiinflammatory; ds.

XX Homo sapiens.  
 OS WO2005054507-A2.  
 PN 16-JUN-2005.  
 PD 03-DEC-2004; 2004WO-GB005078.  
 PF 04-DEC-2003; 2003GB-00028048.  
 PR (UYSH-) UNIV SHEFFIELD.  
 PA Corfe B, Chirakkal H;  
 PI WPI; 2005-435407/44.  
 XX  
 PT Screening for nucleic acid molecules exhibiting altered expression in  
 PT cells grown in the presence of butyrate, and detection of the nucleic  
 PT acid molecules or the encoded polypeptides in diagnosing colorectal  
 PT cancer.  
 XX  
 PS Disclosure; Page 186; 266pp; English.  
 XX  
 CC The invention relates to a method of screening for nucleic acid molecules  
 CC that show altered expression in a first cell sample comprising comparing  
 CC the gene expression profile of the sample with that of a second reference  
 CC sample, where the first sample has been grown in the presence of butyrate  
 CC or a related carbon source from which butyrate is directly or indirectly  
 CC derived, but the reference sample has not. The invention also relates to  
 CC a method of detecting at least one nucleic acid molecule associated with  
 CC the initiation and/or progression of colorectal cancer in an animal,  
 CC comprising providing a biological sample comprising at least one cell to  
 CC be tested, contacting the sample with a ligand (preferably a hybridizing  
 CC nucleic acid molecule) which binds to at least one nucleic acid and  
 CC detecting the presence of at least one molecule in the sample, a method  
 CC of detecting at least one polypeptide associated with the initiation  
 CC and/or progression of colorectal cancer in an animal comprising providing  
 CC a biological sample comprising at least one cell to be tested, contacting  
 CC the sample with at least one ligand that specifically binds at least one  
 CC polypeptide encoded by a nucleic acid or a variant polypeptide comprising  
 CC an amino acid sequence which varies by the addition, deletion or  
 CC substitution of at least one amino acid residue and detecting the  
 CC presence of the polypeptide in the sample, a method of screening for  
 CC agents that modulate the activity of at least one polypeptide encoded by  
 CC a gene associated with the initiation and/or progression of colorectal  
 CC cancer comprising forming a preparation comprising at least one  
 CC polypeptide encoded by a nucleic acid or a variant polypeptide comprising  
 CC at least one addition, deletion or substitution and at least one agent to  
 CC be tested and determining the activity of the agent with respect to  
 CC activity of the polypeptide, and an antibody or its effective binding  
 CC portion identified by the method, for use as a pharmaceutical. The  
 CC methods are useful for screening for nucleic acid molecules that show  
 CC altered expression in a cell sample, and for detecting a nucleic acid and  
 CC a polypeptide respectively, that are associated with the initiation  
 CC and/or progression of colorectal cancer and are useful for detecting or  
 CC monitoring colorectal cancer, especially adenocarcinoma. The methods are  
 CC also useful for screening for agents that modulate the activity of at  
 CC least one polypeptide encoded by a gene associated with the initiation  
 CC and/or progression of cancer, where agents identified by the method are  
 CC useful for treating colorectal cancer. The methods could also be used to  
 CC detect or monitor other conditions such as colitis, Crohn's disease or  
 CC irritable bowel syndrome, as a screening tool for fiber consumption, as  
 CC an assay for colon microflora functionality or for early detection of pre  
 CC -cancerous growth. This sequence represents a human nucleic acid  
 CC identified by the screening method of the invention.  
 XX  
 SQ Sequence 522 BP; 99 A; 171 C; 170 G; 82 T; 0 U; 0 Other;  
 Query Match 94.7%; Score 74.8; DB 14; Length 522;  
 Best Local Similarity 97.4%; Pred. No. 5.7e-14;  
 Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGAATCCCCCTTCAGTCTTTTGAAGAAAGTTCCATGACTCGAATATCTGAATGAAGAA 60  
 |||||  
 Db 82 CGGAATCCCCCTTCAGTCTTTTGAAGAAAGTTCCATGACTCGAATATCTGAATGAAGAA 23  
 |||||  
 Qy 61 AACAAACCGACTCAAAA 78  
 |||||  
 Db 22 AACAAACCGACTCAAAA 5  
 |||||

## RESULT 5

ADP81837  
 ID ADF81837 standard; DNA; 592 BP.

XX AC ADF81837;  
 XX DT 26-FEB-2004 (first entry)  
 XX DE Leukaemia-related DNA sequence #2393.  
 XX KW Cytostatic; Gene therapy; leukaemia; ss.  
 XX OS Unidentified.

XX WO2003039443-A2.  
 XX PD 15-MAY-2003.  
 XX PF 04-NOV-2002; 2002WO-EP012303.  
 XX PR 05-NOV-2001; 2001EP-00126244.  
 XX PR 30-APR-2002; 2002EP-00009758.  
 XX PA (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.  
 XX PA (UFLU-) UNIV LUDWIG MAXIMILIANS.  
 XX PA (HAFE/) HAFERLACH T.  
 XX PA (SCHO/) SCHOCH C.  
 XX PA (KERN/) KERN W.

XX PI Haferlach T, Schoch C, Kern W, Kohlmann A, Schnittger S, Dugas M;  
 XX PI Elis R, Brors B, Mergenthaler S;  
 XX WPI; 2003-505037/47.  
 XX PT Determining the subtype of leukemia cells and whether a patient sample  
 PT contains leukemia cells or other cells, useful for treating leukemia,  
 PT comprises determining the expression profile of a group of markers in a  
 PT patient sample.  
 XX PS Disclosure; SEQ ID NO 2393; 2938pp; English.  
 XX CC The present invention relates to a method (M1) for determining the  
 CC subtype of leukemia cells and whether a patient sample contains  
 CC leukemia cells. The method comprises determining the expression profile  
 CC of a group of markers in a patient sample. The method is useful for  
 CC determining the presence of leukemia cells, its types or subtypes, and  
 CC for the preparation of a medicament for treating leukaemia.  
 XX SQ Sequence 592 BP; 133 A; 173 C; 178 G; 108 T; 0 U; 0 Other;

Query Match 94.7%; Score 74.8; DB 10; Length 592;  
 Best Local Similarity 97.4%; Pred. No. 5.9e-14;  
 Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGAATCCCCCTTCAGTCTTTTGAAGAAAGTTCCATGACTCGAATATCTGAATGAAGAA 60  
 |||||  
 Db 434 CGGAATCCCCCTTCAGTCTTTTGAAGAAAGTTCCATGACTCGAATATCTGAATGAAGAA 493  
 |||||  
 Qy 61 AACAAACCGACTCAAAA 78  
 |||||  
 Db 494 AACAAACCGACTCAAAA 511  
 |||||

## RESULT 6

ADQ86642  
 ID ADQ86642 standard; cDNA; 1985 BP.  
 XX AC ADQ86642;  
 XX DT 07-OCT-2004 (first entry)  
 XX DE Human tumour-associated antigenic target (TAT) cDNA sequence #3516.  
 XX KW human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;  
 KW cancer; cell proliferative disorder; gene; ss.  
 XX OS Homo sapiens.  
 XX PN WO2004060270-A2.  
 XX PD 22-JUL-2004.  
 XX PF 15-OCT-2003; 2003WO-US029126.  
 XX PR 18-OCT-2002; 2002US-0418988P.  
 XX PA (GETH ) GENENTECH INC.  
 XX PA (WUTD/) WU T D.  
 XX PA (ZHOU/) ZHOU Y.  
 XX PI Wu TD, Zhou Y;  
 XX WPI; 2004-534300/51.

XX New nucleic acid molecule and encoded polypeptide, for diagnosing,  
 PT preventing or treating cell proliferative disorders such as cancer.  
 XX Claim 1; SEQ ID NO 3516; 5504pp; English.  
 XX PS The present invention describes an isolated tumour-associated antigenic  
 CC target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide  
 CC sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of  
 CC (a); (c) the complement of (a) or (b); (d) a sequence that has 80%  
 CC sequence identity to (a)-(c); or (e) a sequence that hybridises to (a)-  
 CC (c). Also described: (1) an expression vector comprising the above  
 CC nucleic acid; (2) a host cell comprising the above expression vector; (3)  
 CC a process for producing a polypeptide; (4) an isolated polypeptide  
 CC comprising: (a) an amino acid sequence encoded by any of the above  
 CC nucleotide sequences; (b) an amino acid sequence encoded by the full-  
 CC length coding region of the above nucleotide sequences; or (c) a sequence  
 CC having at least 80% identical to (a) or (b); (5) a chimeric polypeptide  
 CC comprising the above polypeptide fused to a heterologous polypeptide; (6)  
 CC an isolated antibody that binds to the above polypeptide; (7) a process  
 CC for producing the antibody; (8) an isolated oligopeptide that binds to  
 CC the above polypeptide; (9) a tumour-associated antigenic target (TAT)  
 CC binding organic molecule that binds to the above polypeptide; (10) a  
 CC composition of matter comprising the above (chimeric) polypeptide,  
 CC antibody, oligopeptide or TAT binding organic molecule, in combination  
 CC with a carrier; (11) an article of manufacture comprising a container and  
 CC the composition of matter contained within the container; (12) methods of  
 CC inhibiting the growth of a cell that expresses the above protein, where  
 CC the growth of the cell is at least in part dependent upon a growth  
 CC potentiating effect of the above protein; (13) a method of  
 CC therapeutically treating a mammal having a cancerous tumour comprising  
 CC cells that express the above protein; (14) a method of determining the  
 CC presence of a protein in a sample suspected of containing the protein  
 CC described above; (15) methods of diagnosing the presence of a tumour in a  
 CC mammal; (16) a method for treating or preventing a cell proliferative  
 CC disorder associated with increased expression or activity of the above  
 CC protein; and (17) a method of binding an antibody, oligopeptide or  
 CC organic molecule to a cell that expresses the protein described above.  
 CC The TAT sequences have cytostatic activities, and can be used in gene  
 CC therapy. The composition and methods are useful for diagnosing,  
 CC preventing or treating cancer. The composition is also used for preparing  
 CC a medicament for the therapeutic treatment or diagnostic detection of a  
 CC cell proliferative disorder or cancer. The present sequence represents a  
 CC human TAT cDNA sequence from the present invention.

```
XX SQ Sequence 1985 BP; 372 A; 649 C; 594 G; 370 T; 0 U; 0 Other;
XX Query Match 94.7%; Score 74.8; DB 12; Length 1985;
XX Best Local Similarity 97.4%; Pred. No. 7.4e-14;
XX Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 CGGGAATCCCTTCAGTTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
Db 1636 CGGGAATCCCTTCAGTTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1695
Qy 61 AACAAACCGACTCTAAAA 78
Db 1696 AACAAACCGACTCACAAA 1713
RESULT 7
ADQ85525
ID ADQ85525 standard; cDNA; 1985 BP.
AC ADQ85525;
XX 07-OCT-2004 (first entry)
XX Human tumour-associated antigenic target (TAT) cDNA sequence #2339.
XX human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;
XX cancer; cell proliferative disorder; gene; ss.
XX Homo sapiens.
XX WO2004060270-A2.
XX 22-JUL-2004.
XX 15-OCT-2003; 2003WO-US029126.
XX 18-OCT-2002; 2002US-0418988P.
XX (GETH ) GENENTECH INC.
XX (WUTD/) WU T D.
XX (ZHOU/) ZHOU Y.
XX Wu TD, Zhou Y;
XX WPI; 2004-534300/51.
XX New nucleic acid molecule and encoded polypeptide, for diagnosing,
XX preventing or treating cell proliferative disorders such as cancer.
XX Claim 1; SEQ ID NO 2339; 5504pp; English.
XX The present invention describes an isolated tumour-associated antigenic
XX target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide
XX sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of
XX (a); (c) the complement of (a) or (b); (d) a sequence that hybridises to (a)-
XX sequence identity to (a)-(c); or (e) a sequence that hybridises to (a)-
XX nucleic acid; (2) a host cell comprising the above expression vector; (3)
XX a process for producing a polypeptide; (4) an isolated polypeptide
XX comprising: (a) an amino acid sequence encoded by any of the above
XX nucleotide sequences; (b) an amino acid sequence encoded by the full-
XX length coding region of the above nucleotide sequences; or (c) a sequence
XX having at least 80% identical to (a) or (b); (5) a chimeric polypeptide
XX comprising the above polypeptide fused to a heterologous polypeptide; (6)
XX an isolated antibody that binds to the above polypeptide; (7) a process
XX for producing the antibody; (8) an isolated oligopeptide that binds to
XX the above polypeptide; (9) a tumour-associated antigenic target (TAT)
XX binding organic molecule that binds to the above polypeptide; (10) a
XX composition of matter comprising the above (chimeric) polypeptide,
XX antibody, oligopeptide or TAT binding organic molecule, in combination
XX with a carrier; (11) an article of manufacture comprising a container and
XX the composition of matter contained within the container; (12) methods of
```

```
CC inhibiting the growth of a cell that expresses the above protein, where
CC the growth of the cell is at least in part dependent upon a growth
CC potentiating effect of the above protein; (13) a method of
CC therapeutically treating a mammal having a cancerous tumour comprising
CC cells that express the above protein; (14) a method of determining the
CC presence of a protein in a sample suspected of containing the protein in a
CC described above; (15) methods of diagnosing the presence of a tumour in a
CC mammal; (16) a method for treating or preventing a cell proliferative
CC disorder associated with increased expression or activity of the above
CC protein; and (17) a method of binding an antibody, oligopeptide or
CC organic molecule to a cell that expresses the protein described above.
CC The TAT sequences have cytostatic activities, and can be used in gene
CC therapy. The composition and methods are useful for diagnosing,
CC preventing or treating cancer. The composition is also used for preparing
CC a medicament for the therapeutic treatment or diagnostic detection of a
CC cell proliferative disorder or cancer. The present sequence represents a
CC human TAT cDNA sequence from the present invention.
XX
XX SQ Sequence 1985 BP; 372 A; 649 C; 594 G; 370 T; 0 U; 0 Other;
XX Query Match 94.7%; Score 74.8; DB 13; Length 1985;
XX Best Local Similarity 97.4%; Pred. No. 7.4e-14;
XX Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 CGGGAATCCCTTCAGTTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
Db 1636 CGGGAATCCCTTCAGTTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1695
Qy 61 AACAAACCGACTCTAAAA 78
Db 1696 AACAAACCGACTCACAAA 1713
RESULT 8
AAC88152
ID AAC88152 standard; cDNA; 2000 BP.
XX
XX AAC88152;
XX 13-MAR-2001 (first entry)
XX Human K12 nucleotide sequence SEQ ID NO:3.
XX
XX Human; CD7; K12; cognate ligand; cluster of differentiation; cancer;
XX identification; inhibiting T cell proliferation; HIV; infection;
XX activating natural killer cell proliferation; leukaemia; lymphoma;
XX sepsis; graft versus host disease; autoimmune disease; arthritis;
XX multiple sclerosis; rheumatoid arthritis; psoriatic arthritis; lupus;
XX scleroderma; psoriasis; atopic dermatitis; type I diabetes mellitus;
XX Hashimoto's thyroiditis; pernicious anaemia; Addison's disease; uveitis;
XX myasthenia gravis; psoriasis; Guillain-Barre Syndrome; Grave's disease;
XX systemic lupus erythematosus; dermatomyositis; asthma; eczema;
XX atopic dermatitis; contact dermatitis; eczematous dermatitis;
XX seborrhoeic dermatitis; rhinitis; ss.
XX Homo sapiens.
XX WO2000073333-A2.
XX 07-DEC-2000.
XX 26-MAY-2000; 2000WO-US014612.
XX 28-MAY-1999; 99US-0136450P.
XX (IMMV ) IMMUNEX CORP.
XX Lyman SD, Fanslow WC;
XX PI
XX WPI; 2001-061511/07.
XX DR P-PSDB; AAB36658.
XX Stimulating intracellular signaling of CD7 comprises contacting a cell
```







CC obstructive pulmonary disease. The marker gene comprises: (a) a group of  
 CC genes (S1) whose expression levels increase when respiratory epithelial  
 CC cells are stimulated with interleukin-13; or (b) a group of genes (S2)  
 CC whose expression levels decrease when respiratory epithelial cells are  
 CC stimulated with interleukin-13. Also described: (1) a reagent (I) for  
 CC testing for bronchial asthma or chronic obstructive pulmonary disease;  
 CC (2) a kit for screening for a candidate compound for a therapeutic agent  
 CC to treat bronchial asthma or chronic obstructive pulmonary disease; (3)  
 CC an animal model for bronchial asthma or chronic obstructive pulmonary  
 CC disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a  
 CC method for producing an animal model for bronchial asthma or chronic  
 CC obstructive pulmonary disease; (6) a therapeutic agent for bronchial  
 CC asthma or chronic obstructive pulmonary disease, comprising the compound,  
 CC a marker gene or an antisense nucleic acid corresponding to a portion of  
 CC the marker gene, a ribozyme, a polynucleotide that suppresses the  
 CC expression of the gene through an RNAi effect or an antibody recognising  
 CC a protein encoded by a marker gene; and (7) a DNA chip for testing for  
 CC bronchial asthma or a chronic obstructive pulmonary disease, on which a  
 CC probe has been immobilised to assay a marker gene. (I) has respiratory  
 CC and antiasthmatic activities, and can be used in gene therapy. The method  
 CC is useful for testing for or screening for a therapeutic agent for  
 CC bronchial asthma or chronic obstructive pulmonary disease. The present  
 CC sequence is used in the exemplification of the present invention.

XX  
 SQ Sequence 2000 BP; 367 A; 659 C; 602 G; 372 T; 0 U; 0 Other;

Query Match 94.7%; Score 74.8; DB 12; Length 2000;  
 Best Local Similarity 97.4%; Pred. No. 7.5e-14;  
 Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGGGAATCCCTTCAGTCTTTTGAAGTTCATGACTCGAATATCTGAATGAAGAA 60  
 DB 1645 CGGGAATCCCTTCAGTCTTTTGAAGTTCATGACTCGAATATCTGAATGAAGAA 1704

QY 61 AACAAACCGACTCTAAAA 78  
 DB 1705 AACAAACCGACTCACAA 1722

RESULT 12  
 ADR52746  
 ID ADR52746 standard; DNA; 2000 BP.

XX ADR52746;  
 XX 18-NOV-2004 (first entry)  
 XX Drug therapy altered expressed gene #97.  
 XX drug activity monitoring; expression profile; gene expression;  
 KW peripheral blood sample; peripheral blood mononuclear cell; drug therapy;  
 KW CCI-779; immunosuppressant; rapamycin; mammalian target of rapamycin;  
 KW mTOR; ds.  
 XX Homo sapiens.  
 OS  
 XX WO2004072265-A2.  
 XX 26-AUG-2004.  
 XX 11-FEB-2004; 2004WO-US004118.  
 XX 11-FEB-2003; 2003US-0446133P.  
 PR 03-APR-2003; 2003US-0459782P.  
 PR 23-JAN-2004; 2004US-0538246P.  
 XX (AMHP ) WYETH.  
 PA (BURC/) BURCZYNSKI M.  
 PA (TWIN/) TWINE N.  
 PA (DORN/) DORNER A J.  
 PA (TREP/) TREPICCHIO W L.  
 XX Burczynski M, Twine N, Dörner AJ, Trepicchio WL;

XX WPI; 2004-642301/62.

DR Monitoring drug activities in vivo comprises comparing an expression  
 XX profile of a gene in a peripheral blood sample of a patient before and  
 PT after drug therapy.

XX Disclosure; SEQ ID NO 97; 136pp; English.

XX The invention relates to a method of monitoring drug activities in vivo  
 CC by comparing an expression profile of at least one gene in a peripheral  
 CC blood sample of a patient to a reference expression profile of the at  
 CC least one gene, where the at least one gene is differentially expressed  
 CC in peripheral blood mononuclear cells (PBMCs) of patients who have a non-  
 CC blood disease and are subjected to a drug therapy as compared to PBMCs  
 CC isolated from the patient before the drug therapy, and where the patient  
 CC has the non-blood disease and is being treated by the drug therapy. The  
 CC method, kit, and nucleic acid array are useful for monitoring drug  
 CC activities in vivo. The drug is especially CCI-779, an ester analogue of  
 CC the immunosuppressant rapamycin which is a potent inhibitor of the  
 CC mammalian target of rapamycin (mTOR). This sequence represents a gene  
 CC expressed in PMMC altered by the drug therapy. (Note: this sequence does  
 CC no form part of the printed specification but was obtained in electronic  
 CC format from WIPO at ftp.wipo.int/pub/published\_pct\_sequences/).

XX Sequence 2000 BP; 367 A; 659 C; 602 G; 372 T; 0 U; 0 Other;

Query Match 94.7%; Score 74.8; DB 13; Length 2000;  
 Best Local Similarity 97.4%; Pred. No. 7.5e-14;  
 Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGGGAATCCCTTCAGTCTTTTGAAGTTCATGACTCGAATATCTGAATGAAGAA 60  
 DB 1645 CGGGAATCCCTTCAGTCTTTTGAAGTTCATGACTCGAATATCTGAATGAAGAA 1704

QY 61 AACAAACCGACTCTAAAA 78  
 DB 1705 AACAAACCGACTCACAA 1722

RESULT 13  
 ADR24663  
 ID ADP24663 standard; cDNA; 2000 BP.

XX ADP24663;  
 XX 18-NOV-2004 (first entry)  
 XX PRO polypeptide encoding cDNA SEQ ID NO:1841.  
 DE  
 XX ss, gene; PRO; antiinflammatory; antiarthritic; antirheumatic;  
 KW immunosuppressive; osteopathic; antidiabetic; dermatological;  
 KW antipsoriatic; antiallergic; antiasthmatic; hepatotropic; respiratory;  
 KW gene therapy; immune system.  
 XX Unidentified.  
 OS  
 XX WO2004041170-A2.  
 XX 21-MAY-2004.  
 XX 30-OCT-2003; 2003WO-US034312.  
 PF 01-NOV-2002; 2002US-0423394P.  
 PR (GETH ) GENENTECH INC.  
 PA Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WT;  
 PI Wu TD;  
 XX WPI; 2004-419628/39.  
 DR P-PSDB; ADP24664.  
 XX



PT New PRO polypeptides and polynucleotides, useful for treating e.g.  
PT erythematous, rheumatoid arthritis, diabetes mellitus, immune-mediated  
PT renal disease, or demyelinating diseases of the central or peripheral  
XX nervous system.

PS Claim 1; SEQ ID NO 1841; 2940pp; English.

XX  
CC The invention relates to a novel isolated nucleic acid and the PRO  
CC polypeptide encoded by it. A protein of the invention has  
CC antiinflammatory, antiarthritic, antirheumatic, immunosuppressive,  
CC osteopathic, antidiabetic, dermatological, antipsoriatic, antiallergic,  
CC antiasthmatic, hepatotropic, and respiratory activity. A polynucleotide  
CC of the invention may have a use in gene therapy. The PRO polypeptide, its  
CC agonist, antagonist, or antibody that specifically binds to the  
CC polypeptide is useful for treating an immune related disorder such as  
CC systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis,  
CC juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an  
CC idiopathic inflammatory myopathy, Sjogren's syndrome, systemic  
CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune  
CC thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal  
CC disease, a demyelinating disease of the central or peripheral nervous  
CC system, idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome,  
CC a chronic inflammatory demyelinating polyneuropathy, a hepatobiliary  
CC disease, infectious or autoimmune chronic active hepatitis, primary  
CC biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,  
CC inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's  
CC disease, an autoimmune or immune-mediated skin disease, a bullous skin  
CC disease, erythema multiforme, contact dermatitis, psoriasis, an allergic  
CC disease, asthma, allergic rhinitis, atopic dermatitis, food  
CC hypersensitivity, urticaria, an immunologic disease of the lung,  
CC eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity  
CC pneumonitis, a transplantation associated disease, graft rejection or  
CC graft-versus-host disease. The present sequence encodes a PRO protein of  
CC the invention.

XX  
SQ Sequence 2000 BP; 367 A; 659 C; 602 G; 372 T; 0 U; 0 Other;

Query Match 94.7%; Score 74.8; DB 13; Length 2000;  
Best Local Similarity 97.4%; Pred. No. 7.5e-14;  
Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
Qy 1 CGGGAATCCCTTCAGTCTTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60  
Db 1645 CGGGAATCCCTTCAGTCTTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1704  
  
Qy 61 AACAAACCGACTCTCAAAA 78  
Db 1705 AACAAACCGACTCTCAAAA 1722

RESULT 14  
ADU06261  
ID ADU06261 standard; DNA; 2000 BP.

XX  
AC ADU06261;  
XX  
DT 27-JAN-2005 (first entry)  
XX  
DE Novel bronchial cancer-associated human gene SeqID485.  
XX  
KW bronchial cancer; cytostatic; tumour-associated protein;  
XX cancer detection; metastasis; tumour; gene; ds; human.

OS Homo sapiens.  
XX  
PN DE10316701-A1.  
XX  
PD 04-NOV-2004.  
XX  
PF 09-APR-2003; 2003DE-01016701.  
XX  
PR 09-APR-2003; 2003DE-01016701.

PA (HINZ/) HINZMANN B.  
PA (HERM/) HERMANN K.  
PA (CAST/) HEIDEN CASTANOS-VELEZ E.

PI Mennerich D, Bruemendorf T, Heiden E, Hermann K, Kinnemann H;  
PI Li X, Roepcke S, Staub E, Hinzmann B, Rosenthal A, Pflarsky C;

XX  
DR WPI; 2004-786403/78.  
DR P-PSDB; ADU06748.

XX  
PT New nucleic acid, and derived proteins, useful for diagnosis of bronchial  
PT cancer and in screening for therapeutic and diagnostic agents.

XX  
PS Claim 1; SEQ ID NO 485; 1381pp; German.

XX  
CC This invention relates to a novel isolated nucleic acid associated with  
CC bronchial cancer comprising 489 defined sequences given in the  
CC specification. The invention may be useful for the production of  
CC compounds with a cytostatic activity through the inhibition of expression  
CC or activity of tumour-associated proteins. The novel DNA sequences and  
CC the proteins/peptides encoded by them are used for detecting bronchial  
CC cancer or determining the risk of developing it and to screen for  
CC specific binding partners of the DNA or protein sequences, where the  
CC binding partners are potentially useful as agents for treating or  
CC diagnosing bronchial cancer. The DNA or protein sequences can also be  
CC used for prognosis, detection of metastases and for secondary treatment  
CC (of tumours that have been stabilised or are no longer detectable).  
CC Detecting abnormal expression of the DNA sequences provides early  
CC diagnosis of bronchial cancers. The present sequence is that of a novel  
CC bronchial cancer-associated human gene sequence of the invention.

XX  
SQ Sequence 2000 BP; 367 A; 659 C; 602 G; 372 T; 0 U; 0 Other;

Query Match 94.7%; Score 74.8; DB 13; Length 2000;  
Best Local Similarity 97.4%; Pred. No. 7.5e-14;  
Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
Qy 1 CGGGAATCCCTTCAGTCTTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60  
Db 1645 CGGGAATCCCTTCAGTCTTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1704  
  
Qy 61 AACAAACCGACTCTCAAAA 78  
Db 1705 AACAAACCGACTCTCAAAA 1722

RESULT 15  
ADW38555  
ID ADW38555 standard; cDNA; 2000 BP.

XX  
AC ADW38555;  
XX  
DT 24-MAR-2005 (first entry)  
XX  
DE Immunomodulatory gene SECTM1 cDNA.

XX  
KW cytostatic; immunosuppressive; virucide; diagnosis; prognosis;  
KW pharmaceutical; immunotherapy; cancer; cytostatic; neoplasm;  
KW autoimmune disease; immunosuppressive; immune disorder; viral infection;  
XX infection; SECTM1; gene; ss.

OS Homo sapiens.  
XX  
PN WO2005000099-A2.  
XX  
PD 06-JAN-2005.  
XX  
PF 09-JUN-2004; 2004WO-US018461.  
XX  
PR 09-JUN-2003; 2003US-0477291P.  
XX  
PA (GENZ ) GENZYME CORP.

PI Roberts BL;  
XX  
DR MPI; 2005-058046/06.  
DR P-PSDB; ADM38554.  
XX  
PT New isolated blood factor domain polynucleotides having immunomodulatory  
PT activity, useful for aiding in the diagnosis or treating disorders  
PT relating to the immune responses, e.g. cancers, autoimmune diseases, or  
PT viral infections.  
XX  
XX Disclosure; SEQ ID NO 64; 141pp; English.  
PS  
XX The invention describes an isolated polynucleotide (I) encoding a peptide  
CC selected from 21 polynucleotides fully given in the specification, or the  
CC complement of the polynucleotide. Also described are: an isolated peptide  
CC selected from 21 peptides fully given in the specification; a host cell  
CC comprising (i) or the peptide; an antibody that specifically recognizes  
CC and binds the peptide; a composition comprising the host cell, the  
CC peptide, or the antibody, and a pharmaceutical carrier; an immune  
CC effector cell raised in the presence and at the expense of a host cell;  
CC and a method for eliciting a cytolytic response in a subject. Also  
CC disclosed are: a method for monitoring gene expression; a method for  
CC modulating the expression of the immunomodulatory polynucleotides and  
CC expression products; a method for screening for candidate agents that  
CC modulate the expression of the polynucleotide or the expression products  
CC of the polynucleotide; assays for the identification, assessment, and  
CC development of candidate agents capable of modulating the activity of the  
CC polynucleotides or polypeptides; a method for monitoring an immune  
CC response in a subject; and a method for active immunotherapy. The  
CC polynucleotides (e.g., blood factor domains) having immunomodulatory  
CC activity are useful for detecting, diagnosing, prognosing, or monitoring  
CC the progression of a disease. They are useful for aiding in the diagnosis  
CC or treating disorders relating to the immune responses, e.g. cancers,  
CC autoimmune diseases, or viral infections. This sequence encodes  
CC immunomodulatory gene SECTM1.  
XX  
SQ Sequence 2000 BP; 367 A; 659 C; 602 G; 372 T; 0 U; 0 Other;

Query Match 94.7%; Score 74.8; DB 14; Length 2000;  
Best Local Similarity 97.4%; Pred. No. 7.5e-14;  
Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 CGGGAATCCCGCTTCAGTTCTTTGAAAAAGTTCATGACTCGAATATCTGAATGAGAA 60  
Db 1645 CGGGAATCCCGCTTCAGTTCTTTGAAAAAGTTCATGACTCGAATATCTGAATGAGAA 1704  
QY 61 AACAAACCGACTCTAAAA 78  
Db 1705 AACAAACCGACTCAGAA 1722

Search completed: December 7, 2005, 00:03:00  
Job time : 127.194 secs



QY 61 AACAAACCGACTCTATAAA 78  
|||||  
Db 1705 AACAAACCGACTCACAAA 1722

## RESULT 2

US-09-997-165-3  
; Sequence 3, Application US/09997165  
; Patent No. US20020141999A1  
; GENERAL INFORMATION:  
; APPLICANT: Lyman, Stewart D.  
; APPLICANT: Fanelow, William C.  
; TITLE OF INVENTION: LIGAND FOR CD7 AND METHODS OF USE THEREOF  
; FILE REFERENCE: 2913-US  
; CURRENT APPLICATION NUMBER: US/09/997,165  
; CURRENT FILING DATE: 2001-11-27  
; PRIOR APPLICATION NUMBER: PCT/US00/14612  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/136,450  
; PRIOR FILING DATE: 1999-05-28  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 2000  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (119)..(865)  
US-09-997-165-3

Query Match 94.7%; Score 74.8; DB 3; Length 2000;  
Best Local Similarity 97.4%; Pred. No. 1.1e-13;  
Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 CGGGAATCCCCCTTCAGTTCTTTGAAAAAGTTCCATGACTCGAATATCTCGAATGAAGAA 60  
|||||  
Db 1645 CGGGAATCCCCCTTCAGTTCTTTGAAAAAGTTCCATGACTCGAATATCTCGAATGAAGAA 1704  
|||||

QY 61 AACAAACCGACTCTATAAA 78  
|||||  
Db 1705 AACAAACCGACTCACAAA 1722

## RESULT 3

US-10-775-169-97  
; Sequence 97, Application US/10775169  
; Publication No. US2004017543A1  
; GENERAL INFORMATION:  
; APPLICANT: Wyeth  
; APPLICANT: Burczynski, Michael  
; APPLICANT: Twine, Natalie  
; APPLICANT: Dörner, Andrew  
; APPLICANT: Trepicchio, William  
; TITLE OF INVENTION: Method for Monitoring Drug Activities In Vivo  
; FILE REFERENCE: AM101080 (031896-013000)  
; CURRENT APPLICATION NUMBER: US/10/775,169  
; CURRENT FILING DATE: 2004-02-11  
; NUMBER OF SEQ ID NOS: 5278  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 97  
; LENGTH: 2000  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-775-169-97

Query Match 94.7%; Score 74.8; DB 8; Length 2000;  
Best Local Similarity 97.4%; Pred. No. 1.1e-13;  
Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 CGGGAATCCCCCTTCAGTTCTTTGAAAAAGTTCCATGACTCGAATATCTCGAATGAAGAA 60  
|||||

Db 1645 CGGGAATCCCCCTTCAGTTCTTTGAAAAAGTTCCATGACTCGAATATCTCGAATGAAGAA 1704  
QY 61 AACAAACCGACTCTATAAA 78  
|||||  
Db 1705 AACAAACCGACTCACAAA 1722

## RESULT 4

US-10-843-641A-4264  
; Sequence 4264, Application US/10843641A  
; Publication No. US20050064454A1  
; GENERAL INFORMATION:  
; APPLICANT: Avalon Pharmaceuticals, Inc.  
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using  
; FILE REFERENCE: Signature Gene Sets  
; CURRENT APPLICATION NUMBER: US/10/843,641A  
; CURRENT FILING DATE: 2004-05-12  
; PRIOR APPLICATION NUMBER: US/09/873,367  
; PRIOR FILING DATE: 2001-06-05  
; PRIOR APPLICATION NUMBER: US/09/954,531  
; PRIOR FILING DATE: 2001-09-18  
; PRIOR APPLICATION NUMBER: US/09/954,456  
; PRIOR FILING DATE: 2001-09-25  
; PRIOR APPLICATION NUMBER: US/09/962,436  
; PRIOR FILING DATE: 2001-09-25  
; PRIOR APPLICATION NUMBER: US/09/962,832  
; PRIOR FILING DATE: 2001-09-25  
; PRIOR APPLICATION NUMBER: US/09/964,824  
; PRIOR FILING DATE: 2001-09-27  
; PRIOR APPLICATION NUMBER: US/09/967,768  
; PRIOR FILING DATE: 2001-09-28  
; PRIOR APPLICATION NUMBER: US/09/968,007  
; PRIOR FILING DATE: 2001-10-02  
; PRIOR APPLICATION NUMBER: US/09/969,347  
; PRIOR FILING DATE: 2001-10-02  
; PRIOR APPLICATION NUMBER: US/09/969,708  
; PRIOR FILING DATE: 2001-10-03  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 8447  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 4264  
; LENGTH: 2000  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-843-641A-4264

Query Match 94.7%; Score 74.8; DB 9; Length 2000;  
Best Local Similarity 97.4%; Pred. No. 1.1e-13;  
Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 CGGGAATCCCCCTTCAGTTCTTTGAAAAAGTTCCATGACTCGAATATCTCGAATGAAGAA 60  
|||||  
Db 1645 CGGGAATCCCCCTTCAGTTCTTTGAAAAAGTTCCATGACTCGAATATCTCGAATGAAGAA 1704  
|||||

QY 61 AACAAACCGACTCTATAAA 78  
|||||  
Db 1705 AACAAACCGACTCACAAA 1722

## RESULT 5

US-10-631-467-97  
; Sequence 97, Application US/10631467  
; Publication No. US20050208496A1  
; GENERAL INFORMATION:  
; APPLICANT: Genox Research Inc.  
; TITLE OF INVENTION: Method for testing for bronchial asthma, or chronic obstructive p  
; FILE REFERENCE: 3462.1005-000  
; CURRENT APPLICATION NUMBER: US/10/631,467  
; CURRENT FILING DATE: 2003-07-31  
; PRIOR APPLICATION NUMBER: JP 2003-077212  
; PRIOR FILING DATE: 2003-03-20

;; PRIOR APPLICATION NUMBER: JP 2002-229312  
;; PRIOR FILING DATE: 2002-08-06  
;; NUMBER OF SEQ ID NOS: 2086  
;; SOFTWARE: PatentIn version 3.1  
;; SEQ ID NO 97  
;; LENGTH: 2000  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-10-631-467-97

Query Match 94.7%; Score 74.8; DB 9; Length 2000;  
Best Local Similarity 97.4%; Pred. No. 1.1e-13;  
Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60  
Db 1645 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1704  
Qy 61 AACAAACCGACTCTAAAA 78  
Db 1705 AACAAACCGACTCACAA 1722

RESULT 6  
US-10-080-522-2  
;; Sequence 2, Application US/10080522  
;; Publication No. US20030096326A1  
;; GENERAL INFORMATION:  
;; APPLICANT: KAUFMAN, RUSSEL E.  
;; TITLE OF INVENTION: GENE PRODUCT OVER EXPRESSED IN CANCER  
;; NUMBER OF SEQUENCES: 2  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: NIXON & VANDERHYE P.C.  
;; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR  
;; CITY: ARLINGTON  
;; STATE: VIRGINIA  
;; COUNTRY: U.S.A.  
;; ZIP: 22201-4714

;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/10/080,522  
;; FILING DATE: 25-Feb-2002  
;; CLASSIFICATION: <Unknown>  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 09/539,774  
;; FILING DATE: 31-MAR-2000  
;; APPLICATION NUMBER: US 09/210,474  
;; FILING DATE: 14-DEC-1998  
;; APPLICATION NUMBER: US 08/755,559  
;; FILING DATE: 22-NOV-1996

;; ATTORNEY/AGENT INFORMATION:  
;; NAME: WILSON, MARY J.  
;; REGISTRATION NUMBER: 32,955  
;; REFERENCE/DOCKET NUMBER: 1579-645  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (703) 816-4000  
;; TELEFAX: (703) 816-4100  
;; INFORMATION FOR SEQ ID NO: 2:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 2180 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: DNA (genomic)  
;; SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-10-080-522-2

Query Match 94.7%; Score 74.8; DB 5; Length 2180;  
Best Local Similarity 97.4%; Pred. No. 1.1e-13;  
Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60  
Db 1841 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1900  
Qy 61 AACAAACCGACTCTAAAA 78  
Db 1901 AACAAACCGACTCACAA 1918

RESULT 7  
US-10-437-963-1665/c  
;; Sequence 1665, Application US/10437963  
;; Publication No. US2004012343A1  
;; GENERAL INFORMATION:  
;; APPLICANT: La Rosa, Thomas J.  
;; APPLICANT: Kovalic, David K.  
;; APPLICANT: Zhou, Yihua  
;; APPLICANT: Cao, Yongwei  
;; APPLICANT: Wu, Wei  
;; APPLICANT: Boukharov, Andrey A.  
;; APPLICANT: Barbazuk, Brad  
;; APPLICANT: Li, Ping  
;; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
;; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
;; FILE REFERENCE: 38-21(53221)B  
;; CURRENT APPLICATION NUMBER: US/10/437,963  
;; CURRENT FILING DATE: 2003-05-14  
;; NUMBER OF SEQ ID NOS: 204966  
;; SEQ ID NO 1665  
;; LENGTH: 1166  
;; TYPE: DNA  
;; ORGANISM: Oryza sativa  
;; FEATURE:  
;; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_101509C.1  
US-10-437-963-1665

Query Match 38.5%; Score 30.4; DB 7; Length 1166;  
Best Local Similarity 67.2%; Pred. No. 16;  
Matches 43; Conservative 0; Mismatches 21; Indels 0; Gaps 0;  
Qy 13 TTCAGTTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAAACCAACCGACT 72  
Db 1149 TTTTCTTTTATTAAGCAAGCTCAGTCATGATCTGAATTGAAGCAACCAACCGAAT 1090  
Qy 73 CTAA 76  
Db 1089 CTGA 1086

RESULT 8  
US-10-398-221-3432  
;; Sequence 3432, Application US/10398221  
;; Publication No. US20040018514A1  
;; GENERAL INFORMATION:  
;; APPLICANT: KUNST, Frederik  
;; APPLICANT: GLASER, Philippe  
;; TITLE OF INVENTION: Listeria innocua, genome and applications  
;; FILE REFERENCE: 344 702 - US  
;; CURRENT APPLICATION NUMBER: US/10/398,221  
;; CURRENT FILING DATE: 2003-03-27  
;; PRIOR APPLICATION NUMBER: PCT/FR 01/03 061  
;; PRIOR FILING DATE: 2001-10-04  
;; PRIOR APPLICATION NUMBER: FR 00/12 697  
;; PRIOR FILING DATE: 2000-10-04  
;; NUMBER OF SEQ ID NOS: 4025  
;; SOFTWARE: PatentIn version 3.0  
;; SEQ ID NO 3432  
;; LENGTH: 2278  
;; TYPE: DNA

```
; ORGANISM: Listeria monocytogenes 4b
; FEATURES:
; NAME/KEY: misc feature
; LOCATION: (1)..(end)
; OTHER INFORMATION: n can be any nucleotide: a,g,c or t/u
; US-10-398-221-3432

Query Match          37.0%; Score 29.2; DB 7; Length 2278;
Best Local Similarity 62.2%; Pred. No. 48;
Matches 46; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 AATCCCTTCAGTTCCTTTGAAAGTTCCATGACTCGAATATCTGAAATGAAGAAACA 64
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1274 AATCCCTTCATTGCTGAACAACACTTTATTGATTAGATAAAGCATACCAATAAAGA 1333

Qy 65 AACCGACTCTAAAA 78
    |||||
Db 1334 AAGCAGGGATAAAA 1347

RESULT 9
US-09-814-353-17505/c
; Sequence 17505, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17505
; LENGTH: 528
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-814-353-17505

Query Match          36.7%; Score 29; DB 3; Length 528;
Best Local Similarity 61.0%; Pred. No. 36;
Matches 47; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

Qy 2 GGGAAATCCCTTCAGTTCCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAAA 61
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 GGGAAATTCCTATGCAGGCGCTTTAAAAAATAAGTGATTTCGATTTTAAAAAATAAAGAAA 122

Qy 62 ACAACCGACTCTAAAA 78
    |||||
Db 121 AAAAAAATAAAGAAA 105

RESULT 10
US-09-925-065A-787106/c
; Sequence 787106, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.

; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 787107
; LENGTH: 582
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-925-065A-787107

Query Match          36.5%; Score 28.8; DB 4; Length 582;
Best Local Similarity 69.6%; Pred. No. 43;
Matches 39; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 6 ATCCCTTCAGTTCCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAAA 61
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 179 ATGCCTCTCTGATTATTTAAAAAAGTTATTTCATTAGAGTATCAGATTGAAAAAGA 124

RESULT 11
US-09-925-065A-787107/c
; Sequence 787107, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.

; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 787107
; LENGTH: 582
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-925-065A-787107

Query Match          36.5%; Score 28.8; DB 4; Length 582;
Best Local Similarity 69.6%; Pred. No. 43;
Matches 39; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 6 ATCCCTTCAGTTCCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAAA 61
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 179 ATGCCTCTCTGATTATTTAAAAAAGTTATTTCATTAGAGTATCAGATTGAAAAAGA 124

RESULT 12
US-09-925-065A-695404/c
; Sequence 695404, Application US/09925065A
```

```
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 695404
; LENGTH: 1901
; TYPE: DNA
; ORGANISM: Homo sapiens
; ORGANISM: Homo sapiens
US-09-925-065A-695404

Query Match      36.5%; Score 28.8; DB 4; Length 1901;
Best Local Similarity 65.6%; Pred. No. 62;
Matches 42; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy 13 TTCAGTCTTTTGAAGAAAGTTCATGACTCGAATATCTGAAATGAAGAAACAAACCGACT 72
Db 343 TTCATTCTTTGAAATTTTCTTGACTCTATGAATAAAGTGAAGCAACAGAAA 284
```

Qy 73 CTAA 76

Db 283 CAAA 280

RESULT 13

```
US-10-087-192-736
; Sequence 736, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 736
; LENGTH: 31828
; TYPE: DNA
; ORGANISM: Homo sapiens
; ORGANISM: Homo sapiens
US-10-087-192-736
```

```
Query Match      36.5%; Score 28.8; DB 5; Length 31828;
Best Local Similarity 69.6%; Pred. No. 1.4e+02;
Matches 39; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
```

```
Qy 6 ATCCCCCTTCAGTCTTTTGAAGAAAGTTCATGACTCGAATATCTGAAATGAAGAAA 61
Db 20828 ATGCCTCTCTGTATTTTAAAAAAGTTATTTTCATTAGAGATATCAGAAATTGAAAGA 20883
```

RESULT 14

```
US-10-312-841-1/c
; Sequence 1, Application US/10312841
; Publication No. US20030186277A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MHC
; FILE REFERENCE: E01/1208/WO
; CURRENT APPLICATION NUMBER: US/10/312,841
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 2
; SEQ ID NO 1
; LENGTH: 3673778
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (3294164)
US-10-312-841-1
```

```
Query Match      36.2%; Score 28.6; DB 6; Length 3673778;
Best Local Similarity 64.2%; Pred. No. 5.9e+02;
Matches 43; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
```

```
Qy 13 TTCAGTCTTTTGAAGAAAGTTCATGACTCGAATATCTGAAATGAAGAAACAAACCGACT 72
Db 3131340 TCCATATTTTTCAAACTAATACTATATTAATAACGAAAAAAAACCTCCT 3131281
Qy 73 CTAAAAA 79
Db 3131280 TTAATAA 3131274
```

RESULT 15

```
US-09-925-065A-695403/c
; Sequence 695403, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 695403
; LENGTH: 1901
; TYPE: DNA
; ORGANISM: Homo sapiens
; ORGANISM: Homo sapiens
US-09-925-065A-695403
```

```
Query Match      35.9%; Score 28.4; DB 4; Length 1901;
Best Local Similarity 64.1%; Pred. No. 83;
Matches 41; Conservative 1; Mismatches 22; Indels 0; Gaps 0;
```

```
Qy 13 TTCAGTCTTTTGAAGAAAGTTCATGACTCGAATATCTGAAATGAAGAAACAAACCGACT 72
Db 343 TTCATTCTTTGAAATTTTCTTGACTCTATGAATAAAGTGAAGCAACAGAAA 284
Qy 73 CTAA 76
```

Db 283 CAAA 280

Search completed: December 6, 2005, 22:56:59  
Job time : 201.712 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 11:34:11 ; Search time 53.9271 Seconds  
(without alignments)  
455.919 Million cell updates/sec

Title: US-09-980-046B-3

Perfect score: 79

Sequence: 1 cgggaatcccccttcagttc.....aaacaaacgactctaaaaa 79

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3289935 seqs, 155610033 residues

Total number of hits satisfying chosen parameters: 6579870

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications NA New:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*
- 2: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*
- 5: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*
- 7: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*
- 9: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*
- 10: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	27.4	34.7	1131	6 US-10-750-185-51109 Sequence 51109, A
2	26.6	33.7	2995	6 US-10-750-185-35627 Sequence 35627, A
3	26.2	33.2	999	6 US-10-750-185-41906 Sequence 41906, A
4	25.8	32.7	2933	6 US-10-750-185-38312 Sequence 38312, A
5	25.4	32.2	3814	6 US-10-750-185-28705 Sequence 28705, A
6	25.2	31.9	490	7 US-11-108-172-317 Sequence 317, App
7	25.2	31.9	1489	6 US-10-750-185-28550 Sequence 28550, A
8	25	31.6	837	6 US-10-750-185-46504 Sequence 46504, A
9	25	31.6	1872	6 US-10-750-185-52067 Sequence 52067, A
10	25	31.6	2066	6 US-10-523-588-2 Sequence 2, Appli
11	24.8	31.4	600	6 US-10-750-185-3989 Sequence 3989, Ap
12	24.8	31.4	639	6 US-10-793-626-1627 Sequence 1627, Ap
13	24.8	31.4	639	6 US-10-793-626-1965 Sequence 1965, Ap
14	24.8	31.4	1274	6 US-10-750-185-29440 Sequence 29440, A
15	24.8	31.4	1693	6 US-10-750-185-34327 Sequence 34327, A
16	24.8	31.4	3227	6 US-10-793-626-3688 Sequence 3688, Ap
17	24.8	31.4	3227	6 US-10-793-626-3387 Sequence 3387, Ap
18	24.8	31.4	4373	6 US-10-750-185-33226 Sequence 33226, A
19	24.8	31.4	186882	7 US-11-112-908-23 Sequence 23, Appli
20	24.6	31.1	1026	6 US-10-750-185-48087 Sequence 48087, A
21	24.6	31.1	1662	6 US-10-750-185-32969 Sequence 32969, A
22	24.4	30.9	727	6 US-10-750-185-60268 Sequence 60268, A
23	24.4	30.9	1079	6 US-10-750-185-29189 Sequence 29189, A

ALIGNMENTS

RESULT 1

US-10-750-185-51109  
; Sequence 51109, Application US/10750185  
; Publication No. US20050260603A1  
; GENERAL INFORMATION:  
; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MM1100-2  
; CURRENT APPLICATION NUMBER: US/10750,185  
; CURRENT FILING DATE: 2003-12-31  
; PRIOR APPLICATION NUMBER: US 60/437,482  
; PRIOR FILING DATE: 2002-12-31  
; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 51109  
; LENGTH: 1131  
; TYPE: DNA  
; ORGANISM: Bovine 19866880844238  
US-10-750-185-51109

Query Match 34.7%; Score 27.4; DB 6; Length 1131;

Best Local Similarity 65.6%; Pred. No. 3.7;

Matches 40; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 18 TTCTTTGAAAAGTTTCATGACTCGAATATCTGAATGAGAAAACCGACTCTAAA 77

Db 995 TTGCTAAATAATTTTCAATGATGCTTTATTTCTAAAAAAGAAAAACCAATTATGAG 1054

QY 78 A 78

Db 1055 A 1055

RESULT 2

US-10-750-185-35627  
; Sequence 35627, Application US/10750185  
; Publication No. US20050260603A1  
; GENERAL INFORMATION:  
; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.

; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MM1100-2  
; CURRENT APPLICATION NUMBER: US/10/750,185  
; CURRENT FILING DATE: 2003-12-31  
; PRIOR FILING DATE: 2002-12-31  
; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 35627  
; LENGTH: 2995  
; TYPE: DNA  
; ORGANISM: Bovine 19866880955785  
US-10-750-185-35627

Query Match 33.7%; Score 26.6; DB 6; Length 2995;  
Best Local Similarity 63.1%; Pred. No. 8.6; Mismatches 24; Indels 0; Gaps 0;  
Matches 41; Conservative 0;  
Qy 12 CTTGAGTTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAAAAACAACCGAC 71  
Db 2271 CTTGGTATATAGAAATATTCTTTGGCTCAAAAGTGAAAAATAAACAACAAAAAACAAC 2330  
Qy 72 TCTAA 76  
Db 2331 TCAAA 2335

RESULT 3  
US-10-750-185-41906  
; Sequence 41906, Application US/10750185  
; Publication No. US20050260603A1  
; GENERAL INFORMATION:  
; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MM1100-2  
; CURRENT APPLICATION NUMBER: US/10/750,185  
; CURRENT FILING DATE: 2003-12-31  
; PRIOR FILING DATE: 2002-12-31  
; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 41906  
; LENGTH: 999  
; TYPE: DNA  
; ORGANISM: Bovine 19866881144378  
US-10-750-185-41906

Query Match 33.2%; Score 26.2; DB 6; Length 999;  
Best Local Similarity 67.3%; Pred. No. 8.7; Mismatches 18; Indels 0; Gaps 0;  
Matches 37; Conservative 0;  
Qy 20 CTTTGAAGATTTCCATGACTCGAATATCTGAAATGAAGAAAAACAACCGACTCT 74  
Db 567 CTTTAAAGATTTTCAAGATTTAAATCTTTAAATGATTTTAAAAAATGATTTCT 621

RESULT 4  
US-10-750-185-38312  
; Sequence 38312, Application US/10750185  
; Publication No. US20050260603A1  
; GENERAL INFORMATION:  
; APPLICANT: MMI GENOMICS, INC.

; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MM1100-2  
; CURRENT APPLICATION NUMBER: US/10/750,185  
; CURRENT FILING DATE: 2003-12-31  
; PRIOR FILING DATE: 2002-12-31  
; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 38312  
; LENGTH: 2933  
; TYPE: DNA  
; ORGANISM: Bovine 19866880900647  
US-10-750-185-38312

Query Match 32.7%; Score 25.8; DB 6; Length 2933;  
Best Local Similarity 63.9%; Pred. No. 15; Mismatches 22; Indels 0; Gaps 0;  
Matches 39; Conservative 0;  
Qy 12 CTTGAGTTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAAAAACAACCGAC 71  
Db 2065 CTTTATTACTTTGGTCAAGTTCAATTTTCTCAATATAAATGGGAAATAAGATATC 2124  
Qy 72 T 72  
Db 2125 T 2125

RESULT 5  
US-10-750-185-28705  
; Sequence 28705, Application US/10750185  
; Publication No. US20050260603A1  
; GENERAL INFORMATION:  
; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MM1100-2  
; CURRENT APPLICATION NUMBER: US/10/750,185  
; CURRENT FILING DATE: 2003-12-31  
; PRIOR FILING DATE: 2002-12-31  
; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 28705  
; LENGTH: 3814  
; TYPE: DNA  
; ORGANISM: Bovine 198668808046329  
US-10-750-185-28705

Query Match 32.2%; Score 25.4; DB 6; Length 3814;  
Best Local Similarity 64.4%; Pred. No. 22; Mismatches 21; Indels 0; Gaps 0;  
Matches 38; Conservative 0;  
Qy 19 TCTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAAAAACAACCGACTCTAAA 77  
Db 1590 TTTTATAATATCTACAATGAATCCATATTATTAAATCAAGAAAAACAACAGTGAATAA 1648

RESULT 6  
US-11-108-172-317/c  
; Sequence 317, Application US/11108172  
; Publication No. US20050260177A1  
; GENERAL INFORMATION:

```
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secretist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Jiang, Yugu
; APPLICANT: Smith, Carole L.
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Skeiky, Yasir A. W.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedwick Thomas S.
; APPLICANT: Carter, Derrick
; APPLICANT:
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; TITLE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C15
; CURRENT APPLICATION NUMBER: US/11/108,172
; CURRENT FILING DATE: 2005-04-15
; PRIOR APPLICATION NUMBER: US 10/025,380
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: US 09/922,217
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 09/833,263
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US 09/649,811
; PRIOR FILING DATE: 2000-08-28
; PRIOR APPLICATION NUMBER: US 09/609,448
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: US 09/575,251
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 09/519,444
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: US 09/504,629
; PRIOR FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: US 09/480,321
; PRIOR FILING DATE: 2000-01-10
; PRIOR APPLICATION NUMBER: US 09/476,296
; PRIOR FILING DATE: 1999-12-30
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1130
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 317
; LENGTH: 490
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-108-172-317

Query Match 31.9%; Score 25.2; DB 7; Length 490;
Best Local Similarity 71.7%; Pred. No. 15;
Matches 33; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 31 TTCCATGACTCGAATATCTCGAATGAAGAAACAAACCGACTCTAA 76
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 276 TTCTTTCATTATACATCTCGAAGAAAAACAAACCACTCTAA 231

RESULT 7
US-10-750-185-28550/c
; Sequence 28550, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: Denise, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10750185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 317
; LENGTH: 490
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-108-172-317
```

```
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 28550
; LENGTH: 1489
; TYPE: DNA
; ORGANISM: Bovine 19866881078092
US-10-750-185-28550

Query Match 31.9%; Score 25.2; DB 6; Length 1489;
Best Local Similarity 66.7%; Pred. No. 20;
Matches 36; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 8 CCCCTTCAGTCTCTTTCGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAAA 61
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 338 CTCCTTCAATTTATTAATTAATGAAATACATGCTCAATTTAGTTAATAAGNAA 285

RESULT 8
US-10-750-185-46504/c
; Sequence 46504, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: Denise, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10750185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 46504
; LENGTH: 837
; TYPE: DNA
; ORGANISM: Bovine 19866880431603
US-10-750-185-46504

Query Match 31.6%; Score 25; DB 6; Length 837;
Best Local Similarity 58.9%; Pred. No. 20;
Matches 43; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

Qy 7 TCCCTTCAGTCTCTTTCGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAAA 66
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 249 TCTGTGTCATCTCTTTCGAAATTTCTTAATACTAAAGAAAGTTAATAAGCAA 190

Qy 67 CCGACTCTAAAAA 79
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 189 CATCATTTTAAA 177

RESULT 9
US-10-750-185-52067
; Sequence 52067, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: Denise, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
```

```
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 52067
; LENGTH: 1872
; TYPE: DNA
; ORGANISM: Bovine 19866880381128
US-10-750-185-52067

Query Match      31.6%; Score 25; DB 6; Length 1872;
Best Local Similarity 61.5%; Pred. No. 25;
Matches 40; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

Qy 13 TTCAAGTTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAAAAACAAACCGACT 72
Db 1464 TTAGGTCTTTTATATCTCTGCAATTTCTTAAATCTGGACAGAGAAATCTAGGCAAC 1523

Qy 73 CTAAA 77
Db 1524 CCACA 1528

RESULT 10
US-10-523-588-2
; Sequence 2, Application US/10523588
; Publication No. US20050251870A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: CSNK1GS AS MODIFIERS OF THE P21 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX03-057C-US
; CURRENT APPLICATION NUMBER: US/10/523,588
; CURRENT FILING DATE: 2005-02-04
; PRIOR APPLICATION NUMBER: US 60/401,739
; PRIOR FILING DATE: 2002-08-07
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 2066
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-523-588-2

Query Match      31.6%; Score 25; DB 6; Length 2066;
Best Local Similarity 58.9%; Pred. No. 25;
Matches 43; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

Qy 7 TCCCCCTTCAGTTCTTTTGAAGTTCATGACTCGAATATCTGAAATGAAGAAAAACAA 66
Db 1953 TGCACTTTATGTTTATTTTAAAAAACAACAAAAACATGGCAGATGCACAAAAA 2012

Qy 67 CCGACTCTAAAAA 79
Db 2013 AAAAAA 2025

RESULT 11
US-10-750-185-3989
; Sequence 3989, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
```

```
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3989
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Bovine MMBT01900
US-10-750-185-3989

Query Match      31.4%; Score 24.8; DB 6; Length 600;
Best Local Similarity 57.9%; Pred. No. 21;
Matches 44; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

Qy 4 GAATCCCTTCAGTTCTTTTGAAGTTCATGACTCGAATATCTGAAATGAAGAAAAAC 63
Db 310 GAGTCTCCATTCATTGCTCAATAACCAAGATCACACTACTTTCTTAATTCACATAA 369

Qy 64 AAACCGACTCTAAAAA 79
Db 370 GACCTATCTCTAAACA 385

RESULT 12
US-10-793-626-1627
; Sequence 1627, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1627
; LENGTH: 639
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-1627

Query Match      31.4%; Score 24.8; DB 6; Length 639;
Best Local Similarity 63.3%; Pred. No. 22;
Matches 38; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy 20 CTTTGAAGTTCATGACTCGAATATCTGAAATGAAGAAAAACAAACCGACTCTAAAAA 79
Db 152 CTGTGATAGATTTCACTGACTCAAGTTTTCAGTTCAAGTCATCAAGGGGACTGAAAAACA 211

RESULT 13
US-10-793-626-1965
; Sequence 1965, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1965
; LENGTH: 639
```

```
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-1965

Query Match      31.4%; Score 24.8; DB 6; Length 639;
Best Local Similarity 63.3%; Pred. No. 22;
Matches 38; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy 20 CTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAAAAACCAACCGACTCTTAAAAA 79
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 152 CTGATAGATATTTCATCGACTCAAGTGTTCAGTTTCAGTCAATCAAGGGGACTGAAAAACA 211

RESULT 14
US-10-750-185-29440
; Sequence 29440, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 29440
; LENGTH: 1274
; TYPE: DNA
; ORGANISM: Bovine 19866880429041
US-10-750-185-29440

Query Match      31.4%; Score 24.8; DB 6; Length 1274;
Best Local Similarity 57.9%; Pred. No. 26;
Matches 44; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

Qy 4 GAATCCCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAAAAAC 63
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 655 GAGTCTCCATTTCATTGCTCAAATACCAAGAAATCACACTACTTTCTTAAATTCAACATAAA 714

Qy 64 AAACCGACTCTAARAA 79
Db 715 GACTATCTCTAAACA 730

RESULT 15
US-10-750-185-34327
; Sequence 34327, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
```

```
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 34327
; LENGTH: 1693
; TYPE: DNA
; ORGANISM: Bovine 19866881065593
US-10-750-185-34327

Query Match      31.4%; Score 24.8; DB 6; Length 1693;
Best Local Similarity 63.3%; Pred. No. 28;
Matches 38; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy 14 TCAGTCTTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAAAAACCAACCGACTC 73
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1565 TCAGTCTTTTGGAAGAGTGCATCACCTTAAACCGAAGATGAAAAATTTCTAATAAGTCTC 1624

Search completed: December 6, 2005, 23:16:50
Job time : 53.9271 secs
```

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 11:47:47 ; Search time 139.915 Seconds  
(without alignments)  
4239.411 Million cell updates/sec

Title: US-09-980-046B-4

Perfect score: 89

Sequence: 1 cgggaatcccccttcagtc.....aaaaaaaaaaaaaagaaaaa 89

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N Geneseq\_21.\*

- 1: geneseqn1980s.\*
- 2: geneseqn1990s.\*
- 3: geneseqn2000s.\*
- 4: geneseqn2001as.\*
- 5: geneseqn2001bs.\*
- 6: geneseqn2002as.\*
- 7: geneseqn2002bs.\*
- 8: geneseqn2003as.\*
- 9: geneseqn2003bs.\*
- 10: geneseqn2003cs.\*
- 11: geneseqn2003ds.\*
- 12: geneseqn2004as.\*
- 13: geneseqn2004bs.\*
- 14: geneseqn2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	89	100.0	90	4	AAC89735
2	79.4	89.2	1464	3	AAC59772 Human sec
3	71	79.8	79	4	AAC89734
4	70	78.7	428	13	ADU14748
5	70	78.7	522	14	AEA36210 Human nuc
6	70	78.7	592	10	ADF81837 Leukaemia
7	70	78.7	1985	12	ADQ86642 Human tum
8	70	78.7	1985	13	ADQ85525 Human tum
9	70	78.7	2000	4	AAC88152
10	70	78.7	2000	6	ABL65927 Lung can
11	70	78.7	2000	6	ABK84431 Human CDN
12	70	78.7	2000	12	ADJ74845
13	70	78.7	2000	13	ADR52746
14	70	78.7	2000	13	ADP24663
15	70	78.7	2000	13	ADU06261
16	70	78.7	2000	14	ADW38555
17	70	78.7	2180	2	AAV32446
18	58.6	65.8	212	2	AAT26356 Human gen
19	54	60.7	401	10	ADF80286 Leukaemia

#### ALIGNMENTS

RESULT 1

AAC89735

ID AAC89735 standard; cDNA; 90 BP.

XX

AC AAC89735;

XX

DT 12-MAR-2001 (first entry)

XX

DE Human gastrointestinal inflammation-related cDNA, SEQ ID NO: 4.

XX

KW Human; cytostatic; immunomodulator; immunostimulant; vulnerary;

KW anti-inflammatory; neuroprotective; antibacterial; gene therapy;

KW gastrointestinal inflammation; immune system disorder; genetic disorder;

KW cancer; autoimmune disorder; infection; wound healing; ss.

XX

OS Homo sapiens.

XX

PN WO200073324-A2.

XX

PD 07-DEC-2000.

XX

PF 01-JUN-2000; 2000MO-US015191.

XX

PR 01-JUN-1999; 99US-0137058P.

XX

PA (DIGI-) DIGITAL GENE TECHNOLOGIES INC.

XX

PI Youakim A, Dubose RF, Sims JE, Pribyl TM, Hillbush BS, Haseel KW;

XX

DR WPI; 2001-061508/07.

XX

PT New polynucleotides and polypeptides, useful in gene therapy and in diagnosing a pathological condition, e.g. for modulating gene expression in gastrointestinal inflammation, or for treating cancers or genetic disorders.

XX

PS Claim 1; Page 85; 108pp; English.

XX

CC The present sequence is one of a number of isolated human polynucleotides which are useful in gene therapy, and for diagnosing a pathological condition or a susceptibility to it. In particular, the polynucleotides are useful for modulating gene expression in gastrointestinal

CC

CC inflammation. The polynucleotides are useful for chromosome  
 CC identification, controlling gene expression through triple helix  
 CC formation or antisense DNA or RNA, or identifying individuals from minute  
 CC biological samples using DNA-based identification techniques. The  
 CC polynucleotides can also be used as an alternative to restriction  
 CC fragment length polymorphism (RFLP), by determining the actual base-by-  
 CC base DNA sequences of selected portions of an individual's genome. The  
 CC polynucleotides may also be used as molecular weight markers on Southern  
 CC gels, as diagnostic probes for the presence of a specific mRNA, as a  
 CC probe to substract-out known sequences in the process of discovering novel  
 CC polynucleotides, or as an antigen to elicit an immune response. The  
 CC polypeptides are useful in diagnostic procedures to detect a disorder.  
 CC The polynucleotides and polypeptides are useful for preventing, treating  
 CC or ameliorating immune system disorders, genetic disorders, cancers, some  
 CC autoimmune disorders, or infections. The polynucleotides and polypeptides  
 CC are also useful for differentiating, proliferating or attracting cells,  
 CC leading to the regeneration of tissues, especially in wounds or burns.  
 CC The polypeptides and polynucleotides may also be used as a food additive  
 CC or preservative  
 XX  
 SQ Sequence 90 BP; 46 A; 16 C; 12 G; 16 T; 0 U; 0 Other;

Query Match 100.0%; Score 89; DB 4; Length 90;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-10;  
 Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGGAAATCCCTTCAGTCTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60  
 DB 1 CGGGAAATCCCTTCAGTCTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60  
 QY 61 AACAAACCAAAAAAAAAAAAAAAAAAAAAA 89  
 DB 61 AACAAACCAAAAAAAAAAAAAAAAAAAAAA 89

RESULT 2  
 AAC59772  
 ID AAC59772 standard; cDNA; 1464 BP.

AC AAC59772;

XX 26-JAN-2001 (first entry)

DE Human secreted protein gene 35 SEQ ID NO:45.

XX Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;  
 KW antiarthritis; antiproliferative; cytostatic; cardiant; vasotropic;  
 KW cerebroprotective; neurotropic; neuroprotective; antibacterial; virucide;  
 KW fungicide; ophthalmological; gene therapy; autoimmune disease; neoplasm;  
 KW hyperproliferative disorder; cancer; cardiovascular disorder; infection;  
 KW cerebrovascular disorder; angiogenesis; nervous system disorder;  
 KW ocular disorder; wound healing; skin aging; food additive; preservative;  
 KW ss.

XX Homo sapiens.

XX WO200056751-A1.

XX 28-SEP-2000.

XX 09-MAR-2000; 2000WO-US006013.

XX 19-MAR-1999; 99US-0125360P.

XX 11-JUN-1999; 99US-0138626P.

XX 03-DEC-1999; 99US-0168662P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM, Komatsoulis G;

XX WPI; 2000-579482/54.

XX P-PSDB; AAB34611.

PT Isolated nucleic acid molecule encoding a human secreted protein is used  
 PT in preventing, treating or ameliorating a medical condition.  
 XX Claim 1; Page 362; 419pp; English.

XX The polynucleotide sequences given in AAC59738 to AAC59787 encode the  
 CC human secreted proteins given in AAB34577 to AAB34626. AAB34627 to  
 CC AAB34686 represent human secreted polypeptide sequences and proteins  
 CC homologous to them, which are given in the exemplification of the present  
 CC invention. Human secreted proteins have activities based on the tissues  
 CC and cells the genes are expressed in. Example of activities include:  
 CC antiarthritic; immunosuppressive; antirheumatic; antiproliferative;  
 CC cytostatic; cardiant; vasotropic; cerebroprotective; neurotropic;  
 CC neuroprotective; antibacterial; virucide; fungicide; and  
 CC ophthalmological. The polynucleotides and proteins can be used to  
 CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,  
 CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used  
 CC in diagnosing a pathological condition or susceptibility to a  
 CC pathological condition. Disorders which are diagnosed or treated include  
 CC autoimmune diseases, hyperproliferative disorders e.g. neoplasms and  
 CC cancers of the breast or liver, cardiovascular disorders, cerebrovascular  
 CC disorders, angiogenesis, nervous system disorders, infections caused by  
 CC bacteria, viruses and fungi and ocular disorders. The proteins can also  
 CC be used to aid wound healing and epithelial cell proliferation, to  
 CC prevent skin aging due to sunburn, to maintain organs before  
 CC transplantation, for supporting cell culture of primary tissues, to  
 CC regenerate tissues and in chemotaxis. The proteins can also be used as a  
 CC food additive or preservative to increase or decrease storage  
 CC capabilities. AAC59729 to AAC59737 and AAB34576 represent sequences used  
 CC in the exemplification of the present invention

XX Sequence 1464 BP; 309 A; 483 C; 408 G; 264 T; 0 U; 0 Other;

Query Match 89.2%; Score 79.4; DB 3; Length 1464;  
 Best Local Similarity 93.3%; Pred. No. 1.9e-08;  
 Matches 83; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CGGGAAATCCCTTCAGTCTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60  
 DB 1071 CGGGAAATCCCTTCAGTCTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1130  
 QY 61 AACAAACCAAAAAAAAAAAAAAAAAAAAAA 89  
 DB 1131 AACAAACCGACTCACAAAAAAAAAAAAAAAAA 1159

RESULT 3

AAC89734

ID AAC89734 standard; cDNA; 79 BP.

XX AAC89734;

XX 12-MAR-2001 (first entry)

XX Human gastrointestinal inflammation-related cDNA, SEQ ID NO: 3.

XX Human; cytostatic; immunomodulator; immunostimulant; vulnerary;  
 KW anti-inflammatory; neuroprotective; antibacterial; gene therapy;  
 KW gastrointestinal inflammation; immune system disorder; genetic disorder;  
 KW cancer; autoimmune disorder; infection; wound healing; ss.

XX Homo sapiens.

XX WO200073324-A2.

XX 07-DEC-2000.

XX 01-JUN-2000; 2000WO-US015191.

XX 01-JUN-1999; 99US-0137058P.

XX (DIGI-) DIGITAL GENE TECHNOLOGIES INC.



PI Youakim A, Dubose RF, Sims JE, Pribyl TM, Hillbush BS, Haseel KW;  
XX WPI; 2001-061508/07.  
XX  
XX New polynucleotides and polypeptides, useful in gene therapy and in  
PT diagnosing a pathological condition, e.g. for modulating gene expression  
PT in gastrointestinal inflammation, or for treating cancers or genetic  
PT disorders.  
XX  
XX  
XX Claim 1; Page 84; 108pp; English.  
XX  
XX The present sequence is one of a number of isolated human polynucleotides  
CC which are useful in gene therapy, and for diagnosing a pathological  
CC condition or a susceptibility to it. In particular, the polynucleotides  
CC are useful for modulating gene expression in gastrointestinal  
CC inflammation. The polynucleotides are useful for chromosome  
CC identification, controlling gene expression through triple helix  
CC formation or antisense DNA or RNA, or identifying individuals from minute  
CC biological samples using DNA-based identification techniques. The  
CC polynucleotides can also be used as an alternative to restriction  
CC fragment length polymorphism (RFLP), by determining the actual base-by-  
CC base DNA sequences of selected portions of an individual's genome. The  
CC polynucleotides may also be used as molecular weight markers on Southern  
CC gels, as diagnostic probes for the presence of a specific mRNA, as a  
CC probe to subtract-out known sequences in the process of discovering novel  
CC polynucleotides, or as an antigen to elicit an immune response. The  
CC polypeptides are useful in diagnostic procedures to detect a disorder.  
CC The polynucleotides and polypeptides are useful for preventing, treating  
CC or ameliorating immune system disorders, genetic disorders, cancers, some  
CC autoimmune disorders, or infections. The polynucleotides and polypeptides  
CC are also useful for differentiating, proliferating or attracting cells,  
CC leading to the regeneration of tissues, especially in wounds or burns.  
CC The polypeptides and polynucleotides may also be used as a food additive  
CC or preservative  
XX  
XX Sequence 79 BP; 31 A; 18 C; 12 G; 18 T; 0 U; 0 Other;  
SQ

Query Match 79.8%; Score 71; DB 4; Length 79;  
Best Local Similarity 93.7%; Pred. No. 1.3e-06;  
Matches 74; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCTTCAGTCTTTGAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60  
Db |||||  
Qy 61 AACAAACCAAAAAA 79  
Db |||||  
61 AACAAACCGACTCTAAAAA 79

RESULT 4  
ADU14748  
ID ADU14748 standard; DNA; 428 BP.  
XX  
AC ADU14748;  
XX  
XX 27-JAN-2005 (first entry)  
XX  
XX Solid tumour prognosis gene seqid 5187.  
XX  
XX cytostatic; gene therapy; expression profile; solid tumour;  
KW peripheral blood mononuclear cell; PEMC; prognosis; ds.  
XX  
XX Unidentified.  
XX  
XX WO2004097052-A2.  
XX  
XX 11-NOV-2004.  
XX  
XX 29-APR-2004; 2004WO-US013587.  
XX  
XX 29-APR-2003; 2003US-0466067P.  
XX  
XX 23-JAN-2004; 2004US-0538246P.  
XX

XX (AMHP ) WYETH.  
PA (STRA/) STRAHS A.  
XX  
XX Strahs A, Trepicchio WL, Burczynski ME, Twine NC, Slonim DK;  
PI Immerman F, Dornier AJ;  
XX WPI; 2004-804779/79.  
XX  
XX A method, useful for prognosing and treating solid tumor, comprises  
PT comparing an expression profile of a gene expressed in peripheral blood  
PT mononuclear cells to a reference expression profile of a gene.  
XX  
XX Disclosure; Page; 111pp; English.  
XX  
XX The invention describes a method comprising comparing an expression  
CC profile of at least one gene in a peripheral blood sample of a patient to  
CC at least one reference expression profile of the at least one gene, where  
CC the patient has a solid tumor, and each of the gene is differentially  
CC expressed in peripheral blood mononuclear cells (PBMCs) of a first class  
CC of patients as compared to PBMCs of a second class of patients, where  
CC both the first and second classes of patients have the solid tumor, and  
CC each of the first and second classes is a subcluster formed by an  
CC unsupervised clustering analysis of gene expression profiles in PBMCs of  
CC a population of patients who have the solid tumor, and where the  
CC majority of the first class of patients has a first clinical outcome, and  
CC the majority of the second class of patients has a second clinical  
CC outcome. Also described are: a system comprising (i) a memory or a  
CC storage medium including data that represent an expression profile of at  
CC least one gene in a peripheral blood sample of a patient who has a solid  
CC tumor, (ii) at least another storage medium including data that  
CC represent at least one reference expression profile of the gene, (iii) a  
CC program capable of comparing the expression profile to the reference  
CC expression profile, and (iv) a processor capable of executing the  
CC program, where expression levels of the gene in peripheral blood  
CC mononuclear cells of patients who have the solid tumor correlate with  
CC clinical outcomes of the patients; and a nucleic acid or protein array  
CC comprising concentrated probes for solid tumor prognosis genes, where  
CC each of the solid tumor prognosis genes is differentially expressed in  
CC PBMCs of a first class of patients as compared to PBMCs of a second class  
CC of patients, where both the first and second classes of patients have a  
CC solid tumor, and where the first class of patients has a first clinical  
CC outcome, and the second class of patients has a second clinical outcome.  
CC The method, system, and array are useful for prognosing and treating  
CC solid tumors. This sequence represents a solid tumor prognosis gene of  
CC the invention. Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX Sequence 428 BP; 107 A; 108 C; 112 G; 81 T; 0 U; 20 Other;  
SQ

Query Match 78.7%; Score 70; DB 13; Length 428;  
Best Local Similarity 93.6%; Pred. No. 2.1e-06;  
Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCTTCAGTCTTTGAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60  
Db |||||  
100 CGGGAATCCCTTCAGTCTTTGAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 159

Qy 61 AACAAACCAAAAAA 78  
Db |||||  
160 AACAAACCGACTCACAA 177

RESULT 5  
AEA36210/c  
ID AEA36210 standard; DNA; 522 BP.  
XX  
AC AEA36210;  
XX  
XX 25-AUG-2005 (first entry)  
XX  
XX Human nucleic acid sequence #142.  
XX

XX Screening; gene expression; colorectal tumor; colitis; Crohns disease;  
KW irritable bowel syndrome; gastrointestinal disease; cytostatic;  
KW gastrointestinal-gen.; antiinflammatory; ds.  
XX Homo sapiens.  
OS  
XX WO2005054507-A2.  
XX 16-JUN-2005.  
XX  
XX 03-DEC-2004; 2004WO-GB005078.  
XX  
XX 04-DEC-2003; 2003GB-00028048.  
XX (UYSH-) UNIV SHEFFIELD.  
XX Corfe B, Chirakkal H;  
XX WPI; 2005-435407/44.  
XX  
XX Screening for nucleic acid molecules exhibiting altered expression in  
PT cells grown in the presence of butyrate, and detection of the nucleic  
PT acid molecules or the encoded polypeptides in diagnosing colorectal  
PT cancer.  
XX  
XX Disclosure; Page 186; 266pp; English.  
XX  
XX The invention relates to a method of screening for nucleic acid molecules  
CC that show altered expression in a first cell sample comprising comparing  
CC the gene expression profile of the sample with that of a second reference  
CC sample, where the first sample has been grown in the presence of butyrate  
CC or a related carbon source from which butyrate is directly or indirectly  
CC derived, but the reference sample has not. The invention also relates to  
CC a method of detecting at least one nucleic acid molecule associated with  
CC the initiation and/or progression of colorectal cancer in an animal,  
CC comprising providing a biological sample comprising at least one cell to  
CC be tested, contacting the sample with a ligand (preferably a hybridizing  
CC nucleic acid molecule) which binds to at least one nucleic acid and  
CC detecting the presence of at least one molecule in the sample, a method  
CC of detecting at least one polypeptide associated with the initiation  
CC and/or progression of colorectal cancer in an animal comprising providing  
CC a biological sample comprising at least one cell to be tested, contacting  
CC the sample with at least one ligand that specifically binds at least one  
CC polypeptide encoded by a nucleic acid or a variant polypeptide comprising  
CC an amino acid sequence which varies by the addition, deletion or  
CC substitution of at least one amino acid residue and detecting the  
CC presence of the polypeptide in the sample, a method of screening for  
CC agents that modulate the activity of at least one polypeptide encoded by  
CC a gene associated with the initiation and/or progression of colorectal  
CC cancer comprising forming a preparation comprising at least one  
CC polypeptide encoded by a nucleic acid or a variant polypeptide comprising  
CC at least one addition, deletion or substitution and at least one agent to  
CC be tested and determining the activity of the agent with respect to  
CC activity of the polypeptide, and an antibody or its effective binding  
CC portion identified by the method, for use as a pharmaceutical. The  
CC methods are useful for screening for nucleic acid molecules that show  
CC altered expression in a cell sample, and for detecting a nucleic acid and  
CC a polypeptide respectively, that are associated with the initiation  
CC and/or progression of colorectal cancer and are useful for detecting or  
CC monitoring colorectal cancer, especially adenocarcinoma. The methods are  
CC also useful for screening for agents that modulate the activity of at  
CC least one polypeptide encoded by a gene associated with the initiation  
CC and/or progression of cancer, where agents identified by the method are  
CC useful for treating colorectal cancer. The methods could also be used to  
CC detect or monitor other conditions such as colitis, Crohn's disease or  
CC irritable bowel syndrome, as a screening tool for fiber consumption, as  
CC an assay for colon microflora functionality or for early detection of pre  
CC -cancerous growth. This sequence represents a human nucleic acid  
CC identified by the screening method of the invention.  
XX  
XX Sequence 522 BP; 99 A; 171 C; 170 G; 82 T; 0 U; 0 Other;

Query Match 78.7%; Score 70; DB 14; Length 522;  
Best Local Similarity 93.6%; Pred. No. 2.1e-06;  
Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
Qy 1 CGGGAATCCCTTCAGTTCCTTTGAAAAAGTTCGATGCTCGAATATCTGAAATGAAGAA 60  
Db 82 CGGGAATCCCTTCAGTTCCTTTGAAAAAGTTCGATGCTCGAATATCTGAAATGAAGAA 23  
Qy 61 AACAAACCCCAAAAAA 78  
Db 22 AACAAACCGACTCACAAA 5  
RESULT 6  
ADF81837  
ID ADF81837 standard; DNA; 592 BP.  
XX AC  
XX ADF81837;  
XX 26-FEB-2004 (first entry)  
XX Leukaemia-related DNA sequence #2393.  
XX Cytostatic; Gene therapy; leukaemia; ss.  
XX Unidentified.  
XX OS  
XX WO2003039443-A2.  
XX 15-MAY-2003.  
XX 04-NOV-2002; 2002WO-BP012303.  
XX 05-NOV-2001; 2001EP-00126244.  
XX 30-APR-2002; 2002EP-00009758.  
XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.  
XX (UULU-) UNIV LUDWIG MAXIMILIANS.  
XX (HAFE/) HAFERLACH T.  
XX (SCHO/) SCHOCH C.  
XX (KERN/) KERN W.  
XX Haferlach T, Schoch C, Kern W, Kohlmann A, Schmittger S, Dugas M;  
PI Bils R, Brors B, Mergenthaler S;  
XX WPI; 2003-505037/47.  
XX  
XX Determining the subtype of leukemia cells and whether a patient sample  
PT contains leukemia cells or other cells, useful for treating leukemia,  
PT comprises determining the expression profile of a group of markers in a  
PT patient sample.  
XX  
XX Disclosure; SEQ ID NO 2393; 2938pp; English.  
XX  
XX The present invention relates to a method (M1) for determining the  
CC subtype of leukaemia cells and whether a patient sample contains  
CC leukaemia cells. The method comprises determining the expression profile  
CC of a group of markers in a patient sample. The method is useful for  
CC determining the presence of leukaemia cells, its types or subtypes, and  
CC for the preparation of a medicament for treating leukaemia.  
XX  
XX Sequence 592 BP; 133 A; 173 C; 178 G; 108 T; 0 U; 0 Other;  
Qy  
Query Match 78.7%; Score 70; DB 10; Length 592;  
Best Local Similarity 93.6%; Pred. No. 2.1e-06;  
Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
Qy 1 CGGGAATCCCTTCAGTTCCTTTGAAAAAGTTCGATGCTCGAATATCTGAAATGAAGAA 60  
Db 434 CGGGAATCCCTTCAGTTCCTTTGAAAAAGTTCGATGCTCGAATATCTGAAATGAAGAA 493  
Qy 61 AACAAACCCCAAAAAA 78

Db 494 AACAAACCGACTCAGAA 511

RESULT 7

ADQ86642

ID ADQ86642 standard; cdna; 1985 BP.

XX AC ADQ86642;

XX DT 07-OCT-2004 (first entry)

XX DE Human tumour-associated antigenic target (TAT) cdna sequence #3516.

XX human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;

XX cancer; cell proliferative disorder; gene; ss.

XX OS Homo sapiens.

XX FN WO2004060270-A2.

XX PD 22-JUL-2004.

XX PF 15-OCT-2003; 2003WO-US029126.

XX PR 18-OCT-2002; 2002US-0418988P.

XX PA (GETH ) GENENTECH INC.

XX PA (WUTD/) WU T D.

XX PA (ZHOU/) ZHOU Y.

XX PI Wu TD, Zhou Y;

XX DR WPI; 2004-534300/51.

XX New nucleic acid molecule and encoded polypeptide, for diagnosing,

XX preventing or treating cell proliferative disorders such as cancer.

XX Claim 1; SEQ ID NO 3516; 5504pp; English.

XX The present invention describes an isolated tumour-associated antigenic target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of (a); (c) the complement of (a) or (b); (d) a sequence that hybridises to (a) - sequence identity to (a) - (c); or (e) a sequence that hybridises to (a) - (c). Also described: (1) an expression vector comprising the above nucleic acid; (2) a host cell comprising the above expression vector; (3) a process for producing a polypeptide; (4) an isolated polypeptide comprising: (a) an amino acid sequence encoded by any of the above nucleotide sequences; (b) an amino acid sequence encoded by the full-length coding region of the above nucleotide sequences; or (c) a sequence having at least 80% identical to (a) or (b); (5) a chimeric polypeptide comprising the above polypeptide fused to a heterologous polypeptide; (6) an isolated antibody that binds to the above polypeptide; (7) a process for producing the antibody; (8) an isolated oligopeptide that binds to the above polypeptide; (9) a tumour-associated antigenic target (TAT) binding organic molecule that binds to the above polypeptide; (10) a composition of matter comprising the above (chimeric) polypeptide, antibody, oligopeptide or TAT binding organic molecule, in combination with a carrier; (11) an article of manufacture comprising a container and the composition of matter contained within the container; (12) methods of inhibiting the growth of a cell that expresses the above protein, where the growth of the cell is at least in part dependent upon a growth potentiating effect of the above protein; (13) a method of therapeutically treating a mammal having a cancerous tumour comprising cells that express the above protein; (14) a method of determining the presence of a protein in a sample suspected of containing the protein described above; (15) methods of diagnosing the presence of a tumour in a mammal; (16) a method for treating or preventing a cell proliferative disorder associated with increased expression or activity of the above protein; and (17) a method of binding an antibody, oligopeptide or organic molecule to a cell that expresses the protein described above. The TAT sequences have cytostatic activities, and can be used in gene therapy. The composition and methods are useful for diagnosing,

CC preventing or treating cancer. The composition is also used for preparing a medicament for the therapeutic treatment or diagnostic detection of a cell proliferative disorder or cancer. The present sequence represents a human TAT cdna sequence from the present invention.

XX Sequence 1985 BP; 372 A; 649 C; 594 G; 370 T; 0 U; 0 Other;

Query Match 78.7%; Score 70; DB 12; Length 1985;

Best Local Similarity 93.6%; Pred. No. 2.1e-06;

Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Oy 1 CGGGAATCCCTTCAGTCTTTGAAAAAGCTTCATGACTCGAATATCTGAAATGAAGAA 60

Db 1636 CGGGAATCCCTTCAGTCTTTGAAAAAGCTTCATGACTCGAATATCTGAAATGAAGAA 1695

Oy 61 AACAAACCAAAAAA 78

Db 1696 AACAAACCGACTCAGAA 1713

RESULT 8

ADQ85525

ID ADQ85525 standard; cdna; 1985 BP.

XX AC ADQ85525;

XX DT 07-OCT-2004 (first entry)

XX DE Human tumour-associated antigenic target (TAT) cdna sequence #2339.

XX human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;

XX cancer; cell proliferative disorder; gene; ss.

XX OS Homo sapiens.

XX FN WO2004060270-A2.

XX PD 22-JUL-2004.

XX PF 15-OCT-2003; 2003WO-US029126.

XX PR 18-OCT-2002; 2002US-0418988P.

XX PA (GETH ) GENENTECH INC.

XX PA (WUTD/) WU T D.

XX PA (ZHOU/) ZHOU Y.

XX PI Wu TD, Zhou Y;

XX DR WPI; 2004-534300/51.

XX New nucleic acid molecule and encoded polypeptide, for diagnosing,

XX preventing or treating cell proliferative disorders such as cancer.

XX Claim 1; SEQ ID NO 2339; 5504pp; English.

XX The present invention describes an isolated tumour-associated antigenic target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of (a); (c) the complement of (a) or (b); (d) a sequence that hybridises to (a) - sequence identity to (a) - (c); or (e) a sequence that hybridises to (a) - (c). Also described: (1) an expression vector comprising the above nucleic acid; (2) a host cell comprising the above expression vector; (3) a process for producing a polypeptide; (4) an isolated polypeptide comprising: (a) an amino acid sequence encoded by any of the above nucleotide sequences; (b) an amino acid sequence encoded by the full-length coding region of the above nucleotide sequences; or (c) a sequence having at least 80% identical to (a) or (b); (5) a chimeric polypeptide comprising the above polypeptide fused to a heterologous polypeptide; (6) an isolated antibody that binds to the above polypeptide; (7) a process for producing the antibody; (8) an isolated oligopeptide that binds to the above polypeptide; (9) a tumour-associated antigenic target (TAT) binding organic molecule that binds to the above polypeptide; (10) a composition of matter comprising the above (chimeric) polypeptide, antibody, oligopeptide or TAT binding organic molecule, in combination with a carrier; (11) an article of manufacture comprising a container and the composition of matter contained within the container; (12) methods of inhibiting the growth of a cell that expresses the above protein, where the growth of the cell is at least in part dependent upon a growth potentiating effect of the above protein; (13) a method of therapeutically treating a mammal having a cancerous tumour comprising cells that express the above protein; (14) a method of determining the presence of a protein in a sample suspected of containing the protein described above; (15) methods of diagnosing the presence of a tumour in a mammal; (16) a method for treating or preventing a cell proliferative disorder associated with increased expression or activity of the above protein; and (17) a method of binding an antibody, oligopeptide or organic molecule to a cell that expresses the protein described above. The TAT sequences have cytostatic activities, and can be used in gene therapy. The composition and methods are useful for diagnosing,

CC composition of matter comprising the above (chimeric) polypeptide.  
 CC antibody, oligopeptide or TAT binding organic molecule, in combination  
 CC with a carrier; (11) an article of manufacture comprising a container and  
 CC the composition of matter contained within the container; (12) methods of  
 CC inhibiting the growth of a cell that expresses the above protein, where  
 CC the growth of the cell is at least in part dependent upon a growth  
 CC potentiating effect of the above protein; (13) a method of  
 CC therapeutically treating a mammal having a cancerous tumour comprising  
 CC cells that express the above protein; (14) a method of determining the  
 CC presence of a protein in a sample suspected of containing the protein  
 CC described above; (15) methods of diagnosing the presence of a tumour in a  
 CC mammal; (16) a method for treating or preventing a cell proliferative  
 CC disorder associated with increased expression or activity of the above  
 CC protein; and (17) a method of binding an antibody, oligopeptide or  
 CC organic molecule to a cell that expresses the protein described above.  
 CC The TAT sequences have cytostatic activities, and can be used in gene  
 CC therapy. The composition and methods are useful for diagnosing,  
 CC preventing or treating cancer. The composition is also used for preparing  
 CC a medicament for the therapeutic treatment or diagnostic detection of a  
 CC cell proliferative disorder or cancer. The present sequence represents a  
 CC human TAT cDNA sequence from the present invention.

XX Sequence 1985 BP; 372 A; 649 C; 594 G; 370 T; 0 U; 0 Other;  
 SQ Query Match 78.7%; Score 70; DB 13; Length 1985;  
 Best Local Similarity 93.6%; Pred. No. 2.1e-06;  
 Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCTTCAGTCTTTGAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60  
 Db 1636 CGGGAATCCCTTCAGTCTTTGAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1695  
 Qy 61 AACAAACCAAAAAA 78  
 Db 1696 AACAAACCGACTCAGAA 1713

RESULT 9  
 AAC88152  
 ID AAC88152 standard; cDNA; 2000 BP.

XX AC AAC88152;  
 XX 13-MAR-2001 (first entry)  
 XX Human K12 nucleotide sequence SEQ ID NO:3.  
 XX Human; CD7; K12; cognate ligand; cluster of differentiation; cancer;  
 XX identification; inhibiting T cell proliferation; HIV; infection;  
 XX activating natural killer cell proliferation; leukaemia; lymphoma;  
 XX sepsis; graft versus host disease; autoimmune disease; arthritis;  
 XX multiple sclerosis; rheumatoid arthritis; psoriatic arthritis; lupus;  
 XX scleroderma; psoriasis; atopic dermatitis; type I diabetes mellitus;  
 XX Hashimoto's thyroiditis; pernicious anaemia; Addison's disease; uveitis;  
 XX myasthenia gravis; psoriasis; Guillain-Barre Syndrome; Grave's disease;  
 XX systemic lupus erythematosus; dermatomyositis; asthma; eczema;  
 XX atypical dermatitis; contact dermatitis; eczematous dermatitis;  
 XX seborrheic dermatitis; rhinitis; ss.

XX Homo sapiens.  
 XX WO200073333-A2.  
 XX 07-DEC-2000.  
 XX 26-MAY-2000; 2000WO-US014612.  
 XX 28-MAY-1999; 99US-0136450P.  
 XX (IMMUNEX CORP.  
 XX Lyman SD, Fanslow WC;  
 XX

DR WPI; 2001-061511/07.  
 DR P-PSDB; AAB36658.

XX Stimulating intracellular signaling of CD7 comprises contacting a cell  
 PT expressing CD7 with recombinant K12 protein, the cognate ligand of CD7,  
 PT to inhibit T cell proliferation and/or activate natural killer cell  
 PT proliferation.

PS Disclosure; Page 36-38; 42pp; English.

XX The present invention describes a method for stimulating (S) the  
 CC intracellular signaling of CD (cluster of differentiation) 7 comprising  
 CC contacting a cell that expresses CD7 with a recombinant K12 protein (I),  
 CC the cognate ligand of CD7. (S) is useful for inhibiting T cell  
 CC proliferation and/or activating NK (natural killer) cell proliferation of  
 CC and/or inducing NK toxicity in a mammal which involves administration of  
 CC K12 protein. It is also used for treating HIV-1 infection, cancer (T cell  
 CC leukaemia, acute lymphocytic leukaemia, cutaneous T cell lymphoma),  
 CC bacterial and viral infections, mediated by CD7. In the case of treating  
 CC T cell leukaemia the soluble K12 protein is covalently attached to a  
 CC toxin. A disease mediated by CD7 such as sepsis, graft versus host  
 CC disease due to transplantation, autoimmune diseases, multiple sclerosis,  
 CC arthritis, rheumatoid arthritis, psoriatic arthritis, scleroderma, lupus,  
 CC psoriasis, pernicious anaemia, Addison's disease, myasthenia gravis,  
 CC thyroiditis, atopic dermatitis, type I diabetes mellitus, systemic  
 CC uveitis, psoriasis, Guillain-Barre Syndrome, Grave's disease, eczema,  
 CC lupus erythematosus and dermatomyositis, asthma, eczema, atypical  
 CC dermatitis, contact dermatitis, other eczematous dermatitis, seborrheic  
 CC dermatitis, and rhinitis is also treated by administering a K12  
 CC antagonist (neutralising antibody). The present sequence encodes the  
 CC human K12 protein, which is given in the exemplification of the present  
 CC invention

SQ Sequence 2000 BP; 367 A; 659 C; 602 G; 372 T; 0 U; 0 Other;  
 Query Match 78.7%; Score 70; DB 4; Length 2000;  
 Best Local Similarity 93.6%; Pred. No. 2.1e-06;  
 Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCTTCAGTCTTTGAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60  
 Db 1645 CGGGAATCCCTTCAGTCTTTGAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1704  
 Qy 61 AACAAACCAAAAAA 78  
 Db 1705 AACAAACCGACTCAGAA 1722

RESULT 10  
 ABL65927  
 ID ABL65927 standard; DNA; 2000 BP.

XX ABL65927;  
 XX 15-MAY-2002 (first entry)  
 XX Lung cancer related gene sequence SEQ ID NO:4264.  
 XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
 XX stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;  
 XX cytostatic; gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma;  
 XX gene; de.

XX Homo sapiens.  
 XX OS  
 XX WO200194629-A2.  
 XX 13-DEC-2001.  
 XX 30-MAY-2001; 2001WO-US010838.  
 XX 05-JUN-2000; 2000US-0209473P.  
 XX 05-JUN-2000; 2000US-0209531P.



CC agent that modulates the expression of gene(s) from Gs in the tissue. M1  
 CC is useful for detecting GCA; M2 is useful for modulating GA; M3 is useful  
 CC for screening an agent capable of modulating GCA preferably in an  
 CC inflammation in a tissue; M4 is useful for detecting an inflammation  
 CC (especially chronic) in a tissue, an allergic response in a subject,  
 CC exposure of a subject to a pathogen or sterile inflammatory disease (e.g.  
 CC psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis,  
 CC cardiac reperfusion injury, renal reperfusion injury, ARDS, adult  
 CC respiratory distress syndrome, inflammatory bowel disease, Crohn's  
 CC disease, ulcerative colitis, periodontal disease; also bacterial  
 CC infection, viral infection, parasitic infection, protozoal infection,  
 CC fungal infection and M5 is useful for treating one of the above  
 CC conditions. The present sequence represents a gene differentially  
 CC expressed in granulocytes. Note: The sequence data for this patent did  
 CC not form part of the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 CC  
 SQ Sequence 2000 BP; 367 A; 659 C; 602 G; 372 T; 0 U; 0 Other;  
 Query Match 78.7%; Score 70; DB 6; Length 2000;  
 Best Local Similarity 93.6%; Pred. No. 2.1e-06;  
 Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 CGGGAATCCCCCTTCAGTCTTTGAAAAAGTTCATGACTCGAATATCTGAAATGAAGAA 60  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 1645 CGGGAATCCCCCTTCAGTCTTTGAAAAAGTTCATGACTCGAATATCTGAAATGAAGAA 1704  
 QY 61 AACAAACCAAAAAA 78  
 DB ||||||| | | | | |  
 1705 AACAAACCGACTCACAA 1722  
 RESULT 12  
 ADJ74845  
 ID ADJ74845 standard; DNA; 2000 BP.  
 XX  
 AC ADJ74845;  
 XX  
 DT 20-MAY-2004 (first entry)  
 XX  
 DE Marker gene SEQ ID NO:97.  
 XX  
 KW bronchial asthma; chronic obstructive pulmonary disease;  
 KW respiratory epithelial cell; interleukin-13; respiratory; antiasthmatic;  
 KW gene therapy; marker gene; gene; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN EPI394274-A2.  
 XX  
 PD 03-MAR-2004.  
 XX  
 PF 04-AUG-2003; 2003EP-00254857.  
 XX  
 PR 06-AUG-2002; 2002JP-00229312.  
 PR 20-MAR-2003; 2003JP-00077212.  
 XX  
 PA (GENO-) GENOX RES INC.  
 XX  
 PI Ontani N, Sugita Y, Yamaya M, Kubo H, Nagai H, Izuhara K;  
 XX  
 DR WPI; 2004-193155/19.  
 XX  
 PT Testing for bronchial asthma or chronic obstructive pulmonary disease by  
 PT comparing the expression level of a marker gene in a biological sample  
 PT from a subject with the expression level of the gene in a sample from a  
 PT healthy subject.  
 XX  
 PS Claim 1; SEQ ID NO 97; 241pp; English.  
 XX  
 CC The present invention describes a method of testing for bronchial asthma  
 CC or chronic obstructive pulmonary disease. The method comprises

CC determining the expression level of a marker gene in a biological sample  
 CC from a subject, comparing the expression level determined with the  
 CC expression level of the marker gene in a biological sample from a healthy  
 CC subject, and judging whether the subject has bronchial asthma or chronic  
 CC obstructive pulmonary disease. The marker gene comprises: (a) a group of  
 CC genes (S1) whose expression levels increase when respiratory epithelial  
 CC cells are stimulated with interleukin-13; or (b) a group of genes (S2)  
 CC whose expression levels decrease when respiratory epithelial cells are  
 CC stimulated with interleukin-13. Also described: (1) a reagent (I) for  
 CC testing for bronchial asthma or chronic obstructive pulmonary disease;  
 CC (2) a kit for screening for a candidate compound for a therapeutic agent;  
 CC to treat bronchial asthma or chronic obstructive pulmonary disease; (3)  
 CC an animal model for bronchial asthma or chronic obstructive pulmonary  
 CC disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a  
 CC method for producing an animal model for bronchial asthma or chronic  
 CC obstructive pulmonary disease; (6) a therapeutic agent for bronchial  
 CC asthma or chronic obstructive pulmonary disease, comprising the compound,  
 CC a marker gene or an antisense nucleic acid corresponding to a portion of  
 CC the marker gene, a ribozyme, a polynucleotide that suppresses the  
 CC expression of the gene through an RNAi effect or an antibody recognising  
 CC a protein encoded by a marker gene; and (7) a DNA chip for testing for  
 CC bronchial asthma or a chronic obstructive pulmonary disease, on which a  
 CC probe has been immobilised to assay a marker gene. (I) has respiratory  
 CC and antiasthmatic activities, and can be used in gene therapy. The method  
 CC is useful for testing for or screening for a therapeutic agent for  
 CC bronchial asthma or chronic obstructive pulmonary disease. The present  
 CC sequence is used in the exemplification of the present invention.  
 XX  
 SQ Sequence 2000 BP; 367 A; 659 C; 602 G; 372 T; 0 U; 0 Other;  
 Query Match 78.7%; Score 70; DB 12; Length 2000;  
 Best Local Similarity 93.6%; Pred. No. 2.1e-06;  
 Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 CGGGAATCCCCCTTCAGTCTTTGAAAAAGTTCATGACTCGAATATCTGAAATGAAGAA 60  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 1645 CGGGAATCCCCCTTCAGTCTTTGAAAAAGTTCATGACTCGAATATCTGAAATGAAGAA 1704  
 QY 61 AACAAACCAAAAAA 78  
 DB ||||||| | | | | |  
 1705 AACAAACCGACTCACAA 1722  
 RESULT 13  
 ADR52746  
 ID ADR52746 standard; DNA; 2000 BP.  
 XX  
 AC ADR52746;  
 XX  
 DT 18-NOV-2004 (first entry)  
 XX  
 DE Drug therapy altered expressed gene #97.  
 XX  
 KW drug activity monitoring; expression profile; gene expression;  
 KW peripheral blood sample; peripheral blood mononuclear cell; drug therapy;  
 KW CCI-779; immunosuppressant; rapamycin; mammalian target of rapamycin;  
 KW mTOR; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2004072265-A2.  
 XX  
 PD 26-AUG-2004.  
 XX  
 PF 11-FEB-2004; 2004WO-US004118.  
 XX  
 PR 11-FEB-2003; 2003US-0446133P.  
 PR 23-APR-2003; 2003US-0459782P.  
 PR 23-JAN-2004; 2004US-0538246P.  
 XX  
 PA (AMHP ) WYETH.  
 PA (BURC/) BURCZYNSKI M.  
 PA (TWIN/) TWINE N.

PA (DORN/) DORNER A J.  
PA (TREP/) TREPICCHIO W L.  
XX Burczynski M, Twine N, Dornier AJ, Trepicchio WL;  
XX WPI; 2004-642301/62.  
XX Monitoring drug activities in vivo comprises comparing an expression  
XX profile of a gene in a peripheral blood sample of a patient before and  
XX after drug therapy.  
XX Disclosure; SEQ ID NO 97; 136pp; English.  
XX  
XX The invention relates to a method of monitoring drug activities in vivo  
XX by comparing an expression profile of at least one gene in a peripheral  
XX blood sample of a patient to a reference expression profile of the at  
XX least one gene, where the at least one gene is differentially expressed  
XX in peripheral blood mononuclear cells (PBMCs) of patients who have a non-  
XX blood disease and are subjected to a drug therapy as compared to PBMCs  
XX isolated from the patient before the drug therapy, and where the patient  
XX has the non-blood disease and is being treated by the drug therapy. The  
XX method, kit, and nucleic acid array are useful for monitoring drug  
XX activities in vivo. The drug is especially CCI-779, an ester analogue of  
XX the immunosuppressant rapamycin which is a potent inhibitor of the  
XX mammalian target of rapamycin (mTOR). This sequence represents a gene  
XX expressed in PBMC altered by the drug therapy. (Note: this sequence does  
XX no form part of the printed specification but was obtained in electronic  
XX format from WIPO at ftp.wipo.int/pub/published\_pat\_sequences/).  
XX  
XX Sequence 2000 BP; 367 A; 659 C; 602 G; 372 T; 0 U; 0 Other;  
XX  
XX Query Match 78.7%; Score 70; DB 13; Length 2000;  
XX Best Local Similarity 93.6%; Pred. No. 2.1e-06;  
XX Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
XX  
XX 1 CGGGAATCCCCCTTCAGTTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60  
XX |||||  
XX 1645 CGGGAATCCCCCTTCAGTTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1704  
XX  
XX 61 AACAAACCAAAAAA 78  
XX |||||  
XX 1705 AACAAACCGACTCAGAA 1722  
XX  
XX RESULT 14  
XX ADP24663  
XX ID ADP24663 standard; cDNA; 2000 BP.  
XX AC ADP24663;  
XX  
XX 18-NOV-2004 (first entry)  
XX  
XX PRO polypeptide encoding cDNA SEQ ID NO:1841.  
XX  
XX ss: gene; PRO; antiinflammatory; antiarthritic; antirheumatic;  
XX immunosuppressive; osteopathic; antidiabetic; dermatological;  
XX antipsoriatic; antiallergic; antiasthmatic; hepatotropic; respiratory;  
XX gene therapy; immune system.  
XX  
XX Unidentified.  
XX  
XX WO2004041170-A2.  
XX  
XX 21-MAY-2004.  
XX  
XX 30-OCT-2003; 2003WO-US034312.  
XX  
XX 01-NOV-2002; 2002US-0423394P.  
XX  
XX (GETH ) GENENTECH INC.  
XX  
XX Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WI,  
XX Wu TD;

XX WPI; 2004-419628/39.  
XX P-PSDB; ADP24664.  
XX  
XX New PRO polypeptides and polynucleotides, useful for treating e.g.  
XX erythematous, rheumatoid arthritis, diabetes mellitus, immune-mediated  
XX renal disease, or demyelinating diseases of the central or peripheral  
XX nervous system.  
XX  
XX Claim 1; SEQ ID NO 1841; 2940pp; English.  
XX  
XX The invention relates to a novel isolated nucleic acid and the PRO  
XX polypeptide encoded by it. A protein of the invention has  
XX antiinflammatory, antiarthritic, antirheumatic, immunosuppressive,  
XX osteopathic, antidiabetic, dermatological, antipsoriatic, antiallergic,  
XX antiasthmatic, hepatotropic, and respiratory activity. A polynucleotide  
XX of the invention may have a use in gene therapy. The PRO polypeptide, its  
XX agonist, antagonist, or antibody that specifically binds to the  
XX polypeptide is useful for treating an immune related disorder such as  
XX systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis,  
XX juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an  
XX idiopathic inflammatory myopathy, Sjogren's syndrome, systemic  
XX vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune  
XX thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal  
XX disease, a demyelinating disease of the central or peripheral nervous  
XX system, idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome,  
XX a chronic inflammatory demyelinating polyneuropathy, a hepatobiliary  
XX disease, infectious or autoimmune chronic active hepatitis, primary  
XX biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,  
XX inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's  
XX disease, erythema multiforme, contact dermatitis, psoriasis, an allergic  
XX disease, asthma, allergic rhinitis, atopic dermatitis, food  
XX hypersensitivity, urticaria, an immunologic disease of the lung,  
XX eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity  
XX pneumonitis, a transplantation associated disease, graft rejection or  
XX graft-versus-host disease. The present sequence encodes a PRO protein of  
XX the invention.  
XX  
XX Sequence 2000 BP; 367 A; 659 C; 602 G; 372 T; 0 U; 0 Other;  
XX  
XX Query Match 78.7%; Score 70; DB 13; Length 2000;  
XX Best Local Similarity 93.6%; Pred. No. 2.1e-06;  
XX Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
XX  
XX 1 CGGGAATCCCCCTTCAGTTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60  
XX |||||  
XX 1645 CGGGAATCCCCCTTCAGTTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1704  
XX  
XX 61 AACAAACCAAAAAA 78  
XX |||||  
XX 1705 AACAAACCGACTCAGAA 1722  
XX  
XX RESULT 15  
XX ADU06261  
XX ID ADU06261 standard; DNA; 2000 BP.  
XX AC ADU06261;  
XX  
XX 27-JAN-2005 (first entry)  
XX  
XX Novel bronchial cancer-associated human gene SeqID485.  
XX  
XX bronchial cancer; cytostatic; tumour-associated protein;  
XX cancer detection; metastasis; tumour; gene; ds; human.  
XX  
XX Homo sapiens.  
XX  
XX DE10316701-A1.  
XX  
XX 04-NOV-2004.  
XX

PF 09-APR-2003; 2003DE-01016701.  
XX  
PR 09-APR-2003; 2003DE-01016701.  
XX  
PA (HINZ/) HINZMANN B.  
PA (HERM/) HERMANN K.  
PA (CAST/) HEIDEN CASTANOS-VELEZ E.  
XX  
PI Mennerich D, Bruemendorf T, Heiden E, Hermann K, Kinnemann H;  
PI Li X, Roepcke S, Staub E, Hinzmann B, Rosenthal A, Pilarsky C;  
XX  
DR WPI; 2004-786403/78.  
DR P-PSDB; ADU06748.  
XX  
PT New nucleic acid, and derived proteins, useful for diagnosis of bronchial  
PT cancer and in screening for therapeutic and diagnostic agents.  
XX  
PS Claim 1; SEQ ID NO 485; 1381pp; German.  
XX  
CC This invention relates to a novel isolated nucleic acid associated with  
CC bronchial cancer comprising 489 defined sequences given in the  
CC specification. The invention may be useful for the production of  
CC compounds with a cytostatic activity through the inhibition of expression  
CC or activity of tumour-associated proteins. The novel DNA sequences and  
CC the proteins/peptides encoded by them are used for detecting bronchial  
CC cancer or determining the risk of developing it and to screen for  
CC specific binding partners of the DNA or protein sequences, where the  
CC binding partners are potentially useful as agents for treating or  
CC diagnosing bronchial cancer. The DNA or protein sequences can also be  
CC used for prognosis, detection of metastases and for secondary treatment  
CC (of tumours that have been stabilised or are no longer detectable).  
CC Detecting abnormal expression of the DNA sequences provides early  
CC diagnosis of bronchial cancers. The present sequence is that of a novel  
CC bronchial cancer-associated human gene sequence of the invention.  
XX  
SQ Sequence 2000 BP; 367 A; 659 C; 602 G; 372 T; 0 U; 0 Other;  
Query Match 78.7%; Score 70; DB 13; Length 2000;  
Best Local Similarity 93.6%; Pred. No. 2.1e-06;  
Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
Qy 1 CGGGAATCCCGCTTCACTTTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60  
Db 1645 CGGGAATCCCGCTTCACTTTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1704  
Qy 61 AACAAACCAAAAAAAA 78  
Db 1705 AACAAACCGACTCACAAA 1722  
Search completed: December 7, 2005, 00:03:01  
Job time : 140.915 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 11:31:40 ; Search time 221.613 Seconds  
(without alignments)  
3320.997 Million cell updates/sec

Title: US-09-980-046B-4  
Perfect score: 89  
Sequence: 1 cgggaatcccccttcagttc.....aaaaaaaaaaaaaagaaaaa 89

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications NA Main:

- 1: /cgn2\_6/ptodata/1/pubna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/1/pubna/US08\_PUBCOMB.seq.\*
- 3: /cgn2\_6/ptodata/1/pubna/US09A\_PUBCOMB.seq.\*
- 4: /cgn2\_6/ptodata/1/pubna/US09B\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/1/pubna/US10A\_PUBCOMB.seq.\*
- 6: /cgn2\_6/ptodata/1/pubna/US10B\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/1/pubna/US10C\_PUBCOMB.seq.\*
- 8: /cgn2\_6/ptodata/1/pubna/US10D\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/1/pubna/US10E\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/1/pubna/US11\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
1	70	78.7	2000	3	US-09-954-456-1237	Sequence 1237, Ap
2	70	78.7	2000	3	US-09-997-165-3	Sequence 3, Appli
3	70	78.7	2000	8	US-10-775-169-97	Sequence 97, Appl
4	70	78.7	2000	9	US-10-843-641A-4264	Sequence 4264, Ap
5	70	78.7	2000	9	US-10-631-467-97	Sequence 97, Appl
6	70	78.7	2180	5	US-10-080-522-2	Sequence 2, Appli
7	43.4	48.8	1131	8	US-10-723-860-7152	Sequence 7152, Ap
8	43.2	48.5	433	7	US-10-437-963-70672	Sequence 70672, A
9	42.6	47.9	3673778	6	US-10-312-841-2	Sequence 2, Appli
10	42	47.2	397	8	US-10-425-115-123809	Sequence 123809,
11	41.6	46.7	499	8	US-10-425-115-158614	Sequence 158614,
12	41	45.1	670	8	US-10-425-115-31392	Sequence 31392, A
13	40.8	45.8	252	8	US-10-425-115-100475	Sequence 100475,
14	40.8	45.8	11155	6	US-10-311-455-578	Sequence 578, App
15	40.6	45.6	496	7	US-10-437-963-15885	Sequence 15885, A
16	40.6	45.6	1477	6	US-10-410-681-7	Sequence 7, Appli
17	40.4	45.4	1533	7	US-10-723-860-7968	Sequence 7968, Ap
18	40.2	45.2	197	7	US-10-767-701-19171	Sequence 19171, A
19	40.2	45.2	572	8	US-10-425-115-68935	Sequence 68935, A
20	40.2	45.2	1257	8	US-10-425-115-125064	Sequence 125064,
21	40.2	45.2	11745	6	US-10-240-453-205	Sequence 205, App
22	40	44.9	263	8	US-10-425-115-174630	Sequence 174630,
23	40	44.9	298	7	US-10-424-599-120040	Sequence 120040,

ALIGNMENTS

RESULT 1

US-09-954-456-1237  
; Sequence 1237, Application US/09954456  
; Patent NO. US20020115057A1  
; GENERAL INFORMATION:  
; APPLICANT: Young, Paul  
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc  
; TITLE OF INVENTION: Sets  
; FILE REFERENCE: 689290-76  
; CURRENT APPLICATION NUMBER: US/09/954,456  
; PRIOR FILING DATE: 2001-09-18  
; PRIOR APPLICATION NUMBER: US/60/233,617  
; PRIOR FILING DATE: 2000-09-18  
; PRIOR APPLICATION NUMBER: US/60/234,052  
; PRIOR FILING DATE: 2000-09-20  
; PRIOR APPLICATION NUMBER: US/60/234,923  
; PRIOR FILING DATE: 2000-09-25  
; PRIOR APPLICATION NUMBER: US/60/235,134  
; PRIOR FILING DATE: 2000-09-25  
; PRIOR APPLICATION NUMBER: US/60/235,637  
; PRIOR FILING DATE: 2000-09-26  
; PRIOR APPLICATION NUMBER: US/60/235,638  
; PRIOR FILING DATE: 2000-09-26  
; PRIOR APPLICATION NUMBER: US/60/235,711  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: US/60/235,720  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: US/60/235,840  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: US/60/235,863  
; PRIOR FILING DATE: 2000-09-27  
; NUMBER OF SEQ ID NOS: 2276  
; SOFTWARE: Patent in version 3.0  
; SEQ ID NO 1237  
; LENGTH: 2000  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-954-456-1237

Query Match 78.7%; Score 70; DB 3; Length 2000;

Best Local Similarity 93.6%; Pred. No. 1.6e-07;

Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGGGAATCCCCCTTCAGTTCTTTGAAAAGTTCCATGACTGATATCTGAATGAAGAA 60

DB 1645 CGGGAATCCCCCTTCAGTTCTTTGAAAAGTTCCATGACTGATATCTGAATGAAGAA 1704



;; PRIOR APPLICATION NUMBER: JP 2002-229312  
;; PRIOR FILING DATE: 2002-08-06  
;; NUMBER OF SEQ ID NOS: 2086  
;; SOFTWARE: PatentIn version 3.1  
;; SEQ ID NO 97  
;; LENGTH: 2000  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-10-631-467-97

Query Match 78.7%; Score 70; DB 9; Length 2000;  
Best Local Similarity 93.6%; Pred. No. 1.6e-07;  
Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCTTCAGTTCCTTTGAAAAGTTCATGACTCGAATATCTGAAATGAAGAA 60  
Db 1645 CGGGAATCCCTTCAGTTCCTTTGAAAAGTTCATGACTCGAATATCTGAAATGAAGAA 1704

Qy 61 AACAAACCAAAAAA 78  
Db 1705 AACAAACCGACTCACA 1722

## RESULT 6

US-10-080-522-2  
;; Sequence 2, Application US/10080522  
;; Publication No. US20030096326A1  
;; GENERAL INFORMATION:  
;; APPLICANT: KAUFMAN, RUSSEL E.  
;; SLENTZ-KESLER, KIMBERLY  
;; TITLE OF INVENTION: GENE PRODUCT OVER EXPRESSED IN CANCER  
CELLS

;; NUMBER OF SEQUENCES: 2  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: NIXON & VANDERHUYE P.C.  
;; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR  
;; CITY: ARLINGTON  
;; STATE: VIRGINIA  
;; COUNTRY: U.S.A.  
;; ZIP: 22201-4714

;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/10/080,522  
;; FILING DATE: 25-Feb-2002  
;; CLASSIFICATION: <Unknown>

;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 09/539,774  
;; FILING DATE: 31-MAR-2000  
;; APPLICATION NUMBER: US 09/210,474  
;; FILING DATE: 14-DEC-1998  
;; APPLICATION NUMBER: US 08/755,559  
;; FILING DATE: 22-NOV-1996  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: WILSON, MARY J.  
;; REGISTRATION NUMBER: 32,955  
;; REFERENCE/POCKET NUMBER: 1579-645  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (703) 816-4000  
;; TELEFAX: (703) 816-4100

;; INFORMATION FOR SEQ ID NO: 2:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 2180 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: DNA (Genomic)

;; SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-10-080-522-2

Query Match 78.7%; Score 70; DB 5; Length 2180;  
Best Local Similarity 93.6%; Pred. No. 1.7e-07;  
Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCTTCAGTTCCTTTGAAAAGTTCATGACTCGAATATCTGAAATGAAGAA 60  
Db 1841 CGGGAATCCCTTCAGTTCCTTTGAAAAGTTCATGACTCGAATATCTGAAATGAAGAA 1900

Qy 61 AACAAACCAAAAAA 78  
Db 1901 AACAAACCGACTCACA 1918

## RESULT 7

US-10-723-860-7152  
;; Sequence 7152, Application US/10723860  
;; Publication No. US20040253606A1  
;; GENERAL INFORMATION:

;; APPLICANT: Aziz, Natasha  
;; APPLICANT: Ginsburg, Wendy M.  
;; APPLICANT: Zlotnik, Albert  
;; TITLE OF INVENTION: Methods for Diagnosis of Soft Tissue Sarcoma, Compositions &  
;; TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators  
;; FILE REFERENCE: 05882.0193.NPUS01  
;; CURRENT APPLICATION NUMBER: US/10/723,860  
;; PRIOR FILING DATE: 2003-11-26  
;; PRIOR APPLICATION NUMBER: 60/429,739  
;; PRIOR FILING DATE: 2002-11-26  
;; NUMBER OF SEQ ID NOS: 8393  
;; SOFTWARE: PatentIn version 3.2  
;; SEQ ID NO 7152  
;; LENGTH: 1131  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; LOCATION: (12)..(32)  
;; OTHER INFORMATION: n is a, c, g, or t

US-10-723-860-7152

Query Match 48.8%; Score 43.4; DB 8; Length 1131;  
Best Local Similarity 72.7%; Pred. No. 0.46;  
Matches 56; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

Qy 13 TTCAGTTCCTTTGAAAAGTTCATGACTCGAATATCTGAAATGAAGAAACCAACCAAAA 72  
Db 1000 TTAACACTTACAATAATCCAGACTGTCATATTTAAAAAGAAAAA 1059

Qy 73 AAAAAAAGAAAA 89  
Db 1060 AAAAAAAGAAAA 1076

## RESULT 8

US-10-437-963-70672/c  
;; Sequence 70672, Application US/10437963  
;; Publication No. US20040123343A1  
;; GENERAL INFORMATION:

;; APPLICANT: La Rosa, Thomas J.  
;; APPLICANT: Kovalic, David K.  
;; APPLICANT: Zhou, Yihua  
;; APPLICANT: Cao, Yongwei  
;; APPLICANT: Wu, Wei  
;; APPLICANT: Boukharov, Andrey A.  
;; APPLICANT: Barbazuk, Brad  
;; APPLICANT: Li, Ping  
;; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with  
;; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
;; FILE REFERENCE: 38-21(53221)B  
;; CURRENT APPLICATION NUMBER: US/10/437,963  
;; CURRENT FILING DATE: 2003-05-14  
;; NUMBER OF SEQ ID NOS: 204966  
;; SEQ ID NO 70672

```
; LENGTH: 433
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_71220C.1
US-10-437-963-70672

Query Match      48.5%; Score 43.2; DB 7; Length 433;
Best Local Similarity 68.2%; Pred. No. 0.42; Mismatches 0; Gaps 0;
Matches 60; Conservative 0; Indels 28; Indels 0; Gaps 0;

Qy  2 GGGAAATCCCTTCAGTTCCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAAA 61
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  93 GGGAAATCCCTTCCTTTCTTTTAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 34

Qy  62 ACAAACCAAAAAAAAAAAAAAAAAAGAAAAA 89
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  33 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 6

RESULT 9
US-10-312-841-2/c
; Sequence 2, Application US/10312841
; Publication No. US20030186277A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MHC
; FILE REFERENCE: E01/1208/WO
; CURRENT APPLICATION NUMBER: US/10/312,841
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 2
; SEQ ID NO 2
; LENGTH: 3673778
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (379615)
US-10-312-841-2

Query Match      47.9%; Score 42.6; DB 6; Length 3673778;
Best Local Similarity 78.5%; Pred. No. 4; Mismatches 14; Indels 0; Gaps 0;
Matches 51; Conservative 0; Indels 14; Indels 0; Gaps 0;

Qy  25 AAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAAACCAACCAAAAAAAAAAAAAAG 84
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  775209 AAAAAATTAAAACTTCATATTTAAAAATAAAAAAAAAAAACAAACAAAAAAAAA 775150

Qy  85 AAAAA 89
    |||||
Db  775149 AAAAA 775145

RESULT 10
US-10-425-115-123809/c
; Sequence 123809, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 123809
; LENGTH: 397
; TYPE: DNA
```

```
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_44396C.1
US-10-425-115-123809

Query Match      47.2%; Score 42; DB 8; Length 397;
Best Local Similarity 69.5%; Pred. No. 0.82; Mismatches 25; Indels 0; Gaps 0;
Matches 57; Conservative 0; Indels 25; Indels 0; Gaps 0;

Qy  8 CCCCTTCAGTTCCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAAACAAAC 67
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  139 CCCCATTAAGTTATTGAAGCCGTGGCTTCCTTGGGTTTCCCAAAAAAAAAAAAAA 80

Qy  68 CAAAAAATAAAAAAAAAAGAAAAA 89
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  79 AAAAAAAAAAAAAAAAAAAAAA 58

RESULT 11
US-10-425-115-158614/c
; Sequence 158614, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 158614
; LENGTH: 499
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_76234C.1
US-10-425-115-158614

Query Match      46.7%; Score 41.6; DB 8; Length 499;
Best Local Similarity 67.0%; Pred. No. 1.1; Mismatches 29; Indels 0; Gaps 0;
Matches 59; Conservative 0; Indels 29; Indels 0; Gaps 0;

Qy  2 GGAATCCCTTCAGTTCCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAAA 61
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  133 GGGGGGCCCCCCCCAAGGAGTTTGNAAAAAACCCATGTTCCGAAAAAAAAAAAAA 74

Qy  62 ACAAACCAAAAAAAAAAAAAAAAAAGAAAAA 89
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  73 AAAAAAAAAAAAAAAAAAAAAA 46

RESULT 12
US-10-425-115-31392/c
; Sequence 31392, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 31392
; LENGTH: 670
; TYPE: DNA
; ORGANISM: Zea mays
```

```
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_128637C.1
US-10-425-115-31392

Query Match      46.1%; Score 41; DB 8; Length 670;
Best Local Similarity 66.3%; Pred. No. 1.6;
Matches 59; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCTTCAGTTCCTTGAAGAAAGTTCATGACTCGAATATCTGAAATGAAGAA 60
Db 154 CGGGGCCCCCTTTTTCCTGGGAATTTGGGTTTGTTCACAAAAA 95

Qy 61 AACAAACCAAAAAAAGAAAAA 89
Db 94 AAAAAA 66

RESULT 13
US-10-425-115-100475/c
; Sequence 100475, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 100475
; LENGTH: 252
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(252)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_23146C.1
US-10-425-115-100475

Query Match      45.8%; Score 40.8; DB 8; Length 252;
Best Local Similarity 65.1%; Pred. No. 1.5;
Matches 54; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

Qy 7 TCCCTTCAGTTCCTTGAAGAAAGTTCATGACTCGAATATCTGAAATGAAGAAACAAA 66
Db 105 TCCTCCCTATTATAAGAAATGTTTCATTGTTTAAANNNGAANAANAANAANA 46

Qy 67 CCAAAAAAAGAAAAA 89
Db 45 AAAAAA 23

RESULT 14
US-10-311-455-578/c
; Sequence 578, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: FIEFENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ
; TITLE OF INVENTION: cytosine methylation
; FILE REFERENCE: 5013,1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
```

```
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 578
; LENGTH: 11155
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-578

Query Match      45.8%; Score 40.8; DB 6; Length 11155;
Best Local Similarity 67.9%; Pred. No. 3.3;
Matches 57; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

Qy 6 ATCCCCCTTCAGTTCCTTGAAGAAAGTTCATGACTCGAATATCTGAAATGAAGAAACAA 65
Db 1913 ATAATCCATATTTATTACAAAAAATCCGAAACTTCAATATTTAAAAA 1854

Qy 66 ACCAAAAAAGAAAAA 89
Db 1853 ACATAAAAAA 1830

RESULT 15
US-10-437-963-15885/c
; Sequence 15885, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 15885
; LENGTH: 496
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(496)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_21687C.1
US-10-437-963-15885

Query Match      45.6%; Score 40.6; DB 7; Length 496;
Best Local Similarity 65.9%; Pred. No. 1.9;
Matches 58; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

Qy 2 GGGAAATCCCTTCAGTTCCTTGAAGAAAGTTCATGACTCGAATATCTGAAATGAAGAA 61
Db 127 GGGGGCCCCCTTCGGGCTTTGNAANAANAANAANAANAANAANAANA 68

Qy 62 ACAAACCAAAAAAAGAAAAA 89
Db 67 AAAAAA 40

Search completed: December 6, 2005, 22:57:04
Job time : 226.613 secs
```

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 21:43:57 ; Search time 1015.87 Seconds  
(without alignments)  
4099.008 Million cell updates/sec

Title: US-09-980-046B-4  
Perfect score: 89  
Sequence: 1 cgggaatcccttcagtc.....aaaaaaaaaaaaaagaaaaa 89

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST: \*  
1: gb\_est1: \*  
2: gb\_est2: \*  
3: gb\_est3: \*  
4: gb\_hc: \*  
5: gb\_est4: \*  
6: gb\_est5: \*  
7: gb\_est6: \*  
8: gb\_est7: \*  
9: gb\_ges1: \*  
10: gb\_ges2: \*  
11: gb\_ges3: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	84.2	94.6	693	CB048392	CB048392 NISC_gj04
C 2	82.8	93.0	519	AW014948	AW014948 UI-H-B10-
C 3	79.4	89.2	338	CK903980	CK903980 ie69h03.y
C 4	79.4	89.2	633	BQ182988	BQ182988 UI-H-ED1-
C 5	79.4	89.2	684	CB241372	CB241372 UI-CF-FN0
C 6	79.4	89.2	691	CD364677	CD364677 UI-H-F72-
C 7	75.4	84.7	313	BM053069	BM053069 ie69h03.y
C 8	73	82.0	292	BF431778	BF431778 nab50c04.
C 9	73	82.0	310	AA903087	AA903087 ok46d04.s
C 10	73	82.0	339	CK903979	CK903979 ie69h03.x
C 11	73	82.0	475	AI927624	AI927624 wo85a02.x
C 12	72	80.9	716	BF002331	BF002331 7h22b08.x
C 13	71.6	80.4	445	AI480347	AI480347 tm51c09.x
C 14	71.4	80.2	687	BU689727	BU689727 UI-CF-FN0
C 15	70	78.7	109	AA826896	AA826896 nr86c10.s
C 16	70	78.7	180	AW874230	AW874230 hq86h05.x
C 17	70	78.7	272	BF002104	BF002104 7h02g04.x
C 18	70	78.7	332	AI621135	AI621135 tu49d01.x
C 19	70	78.7	376	AI281889	AI281889 qt8xall.x
C 20	70	78.7	417	AW002382	AW002382 wu61c03.x
C 21	70	78.7	418	AW594044	AW594044 hg35d12.x
C 22	70	78.7	420	AW195936	AW195936 xl85a06.x

C 23	70	78.7	429	1	AW207187	AW207187 UI-H-B11-
C 24	70	78.7	432	1	AW057702	AW057702 wx02c07.x
C 25	70	78.7	440	1	AI218726	AI218726 oo07a10.x
C 26	70	78.7	442	1	AI652267	AI652267 wb28d06.x
C 27	70	78.7	445	1	AW058447	AW058447 wx21a06.x
C 28	70	78.7	453	1	AI805092	AI805092 tu33g10.x
C 29	70	78.7	479	5	EX108872	EX108872 BX108872
C 30	70	78.7	522	2	BF939675	BF939675 nac79g07.
C 31	70	78.7	1633	4	CR595956	CR595956 full-leng
C 32	69	77.5	429	1	AW081554	AW081554 xc42g10.x
C 33	69	77.5	456	1	AI979001	AI979001 wr69a10.x
C 34	68.4	76.9	408	2	BE677472	BE677472 7d84g07.x
C 35	68.4	76.9	413	1	AI433486	AI433486 ti53b01.x
C 36	68.4	76.9	413	1	AI433494	AI433494 ti53c03.x
C 37	68.4	76.9	1650	4	CR594943	CR594943 full-leng
C 38	67.6	76.0	322	5	BX378127	BX378127 BX378127
C 39	66.8	75.1	157	1	AI277084	AI277084 ql40h08.x
C 40	66.8	75.1	350	1	AW974325	AW974325 EST386428
C 41	64.2	72.1	158	1	AI886304	AI886304 wm67a04.x
C 42	62.2	69.9	799	5	BX380143	BX380143 BX380143
C 43	61.6	69.2	1064	5	BX335528	BX335528 BX335528
C 44	58.6	65.8	209	5	C01591	C01591 HUMGS000859
C 45	58	65.2	1700	4	CR614987	CR614987 full-leng

## ALIGNMENTS

RESULT 1  
CB048392/c  
LOCUS  
DEFINITION NISC\_gj04e10.x1 NCI\_CGAP\_Pr28 Homo sapiens linear EST 17-JAN-2003  
3', mRNA sequence.  
ACCESSION CB048392  
VERSION CB048392.1 GI:27786679  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 693)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-re@mail.nih.gov  
CDNA Library Preparation:  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL  
DNA Sequencing by: National Institutes of Health Intramural  
Sequencing Center (NISC)  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
[info@image.llnl.gov](mailto:info@image.llnl.gov)  
Plate: L1AM8006 row: J column: 20  
Seq primer: -21M13 forward primer (ABI).  
Location/Qualifiers  
1. .693  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3270643"  
/sex="male"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_lib="NCI\_CGAP\_Pr28"  
/note="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia)  
with a modified polylinker; Plasmid DNA from the  
normalized library NCI\_CGAP\_Pr22 was prepared, and ss  
circles were made in vitro. Following HAP purification,  
this DNA was used as tracer in a subtractive hybridization  
reaction. The driver was PCR-amplified cDNAs from a pool

of 5,000 clones made from the same library (cloneIDs 985608-986759, 1101192-1101959, and 1217928-1220615). Subtraction by Bento Soares and M. Fatima Bonaldo.

## ORIGIN

Query Match 94.6%; Score 84.2; DB 6; Length 693;  
Best Local Similarity 96.6%; Pred. No. 7.4e-09;  
Matches 86; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCATGACGCGAATATCTGAAATGAAGAA 60  
|||||  
Db 92 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCATGACGCGAATATCTGAAATGAAGAA 33  
|||||

Qy 61 AACAAACCAAAAAAAGAAAAAAGAAAAA 89  
|||||  
Db 32 AACAAACCGAAAAAAGAAAAAAGAAAAA 4  
|||||

## RESULT 2

AW014948/c  
LOCUS AW014948 519 bp mRNA linear EST 10-SEP-1999  
DEFINITION UI-H-BIO-aaf-h-07-0-UI.s1 NCI\_CGAP\_Sub1 Homo sapiens cDNA clone  
IMAGE:2709349 3', mRNA sequence.

ACCESSION AW014948  
VERSION AW014948.1 GI:5863705

## KEYWORDS

SOURCE Homo sapiens (human)

## ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.

1 (bases 1 to 519)

NCI\_CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: [cgapb@mail.nih.gov](mailto:cgapb@mail.nih.gov)

The sequence contained an oligo-dT track that was present in the  
oligonucleotide that was used to prime the synthesis of first  
strand cDNA and therefore this may represent a bonafide poly A  
tail. cDNA library Preparation: M.B. Soares Lab Clone distribution:  
NCI-CGAP clone distribution information can be found through the  
I.M.A.G.E. Consortium/LLNL at:

[www.bio.llnl.gov/bbrp/image/image.html](http://www.bio.llnl.gov/bbrp/image/image.html) The following repetitive

elements were found in this cDNA sequence: 271-336,

>GC richLow complexity

Seq primer: M13 Forward

POLYA=Yes.

## FEATURES

source

Location/Qualifiers

1..519  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2709349"  
/lab\_host="DH10B (Life Technologies)"  
/clone\_lib="NCI\_CGAP Sub1"  
/note="Vector: p773D-Pac (Pharmacia) with a modified  
polylinker; Site 1: Not I; Site 2: Eco RI; The  
NCI\_CGAP Sub1 library is a subtracted library derived from  
BI\_BI Constitutes a mixture of 21 normalized or  
subtracted NCI\_CGAP libraries: NCI\_CGAP\_C04,  
NCI\_CGAP\_Pr28, NCI\_CGAP\_Pr28, NCI\_CGAP\_C010,  
NCI\_CGAP\_C016, NCI\_CGAP\_Kid3, NCI\_CGAP\_Kid12,  
NCI\_CGAP\_Kid3, NCI\_CGAP\_Kid11, NCI\_CGAP\_Lym2,  
NCI\_CGAP\_Br2, NCI\_CGAP\_C08, NCI\_CGAP\_C04, NCI\_CGAP\_Le12,  
NCI\_CGAP\_Brn23, NCI\_CGAP\_Lu5, NCI\_CGAP\_Lu24,  
NCI\_CGAP\_Lu19, NCI\_CGAP\_GC4, NCI\_CGAP\_GC6, NCI\_CGAP\_Brn25.  
These 21 libraries were pooled and a single-stranded DNA  
preparation of the resulting mixture was used as a tracer  
in a subtractive hybridization with a driver whose  
composition is detailed below: NCI\_CGAP\_Kid3 pool 1 LLAM  
3334-3337, 3682-3683, 3798-3803 (IMAGE CloneIDs

## ORIGIN

Query Match 93.0%; Score 82.8; DB 1; Length 519;  
Best Local Similarity 97.7%; Pred. No. 1.5e-08;  
Matches 84; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCATGACGCGAATATCTGAAATGAAGAA 60  
|||||

Db 86 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCATGACGCGAATATCTGAAATGAAGAA 27  
|||||

Qy 61 AACAAACCAAAAAAAGAAAAAAGAAAAA 86  
|||||

Db 26 AACAAACCGAAAAAAGAAAAAAGAAAAA 1  
|||||

## RESULT 3

CK903980 338 bp mRNA linear EST 11-MAR-2004  
DEFINITION te69h03.y5 Melton Normalized Human Islet 4 M4-HIS 1 Homo sapiens  
cDNA clone IMAGE:5672308 5', mRNA sequence.

ACCESSION CK903980

VERSION CK903980.1 GI:45365511

## KEYWORDS

SOURCE Homo sapiens (human)

## ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.

1 (bases 1 to 338)

Melton.D., Meadows.A., Clifton.S., Hillier.L., Marra.M., Pape.D.,  
Wyllie.T., Martin.J., Blistain.A., Schmitt.A., Theising.B.,  
Ritter.E., Ronko.I., Bennett.J., Cardenas.M., Gibbons.M.,  
McCann.R., Cole.R., Teagareishvili.R., Williams.T., Jackson.Y. and  
Bowers.Y.

## REFERENCE

WashU-Harvard Pancreas EST Project  
Unpublished (2000)

## AUTHORS

Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue

## JOURNAL

## COMMENT

Endocrine Pancreas Consortium  
Harvard University, Howard Hughes Medical Institute  
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,  
MA 02138

Tel: 617-495-1812

Fax: 617-495-8557

Email: [dmelton@biohp.harvard.edu](mailto:dmelton@biohp.harvard.edu)

This read is a 5' RESEQUENCE of a previously sequenced pancreas  
clone

This read has been verified (found to hit its original self in the  
correct orientation)

Putative full length read

vector to vector length is

Seq primer: -40RP from Gibco.

Location/Qualifiers

1..338

source



```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5672308"
/sex="Both"
/tissue_type="Islets of Langerhans"
/dev_stage="Adult"
/lab_host="DH10B"
/clone_lib="Melton Normalized Human Islet 4 N4-HIS 1"
/notes="Organ: Pancreas; Vector: pSPORT1; Site 1: Not 1; Site 2: Sal 1; Starting library constructed using SuperScript Plasmid Library kit (Life Technologies). cDNA made by oligo-dT priming. Size-selected by column fractionation; average insert size 1.08 kb. Library was amplified once on solid support and plasmid DNA from library was prepared. The library DNA was normalized by method #4 from Bonaldo, Lennon, and Soares 1996 Genome Research 6:791-806; 0.5 microgram single-stranded library plasmid DNA was mixed with 5 micrograms PCR product representing library inserts and hybridized to an Ecot of 20. Single-stranded (unhybridized) plasmids were isolated by hydroxyapatite chromatography and used to make this library."

```

## ORIGIN

```

Query Match      89.2%; Score 79.4; DB 7; Length 338;
Best Local Similarity 93.3%; Pred. No. 8.8e-08;
Matches 83; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CGGGAATCCCCCTTCAGTTCCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
   |||||
Db 247 CGGGAATCCCCCTTCAGTTCCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 306
   |||||

QY 61 AACAAACCCAAAAAAGAAAAAGAAAAA 89
   |||||
Db 307 AACAAACCGACTCACAAAAAAGAAAAA 335
   |||||

```

## RESULT 4

```

BQ182988/c
LOCUS      BQ182988      633 bp      mRNA      linear      EST 15-JUL-2003
DEFINITION UI-H-ED1-avj1-1-19-0-UI.s2 NCI CGAP ED1 Homo sapiens cDNA clone
IMAGE:5839650 3', mRNA sequence.
ACCESSION  BQ182988
VERSION    BQ182988.1 GI:20358538
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homnidae; Homo.
REFERENCE  1 (bases 1 to 633)
            NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
            National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
            Unpublished (1997)
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Jose Mercuende
            cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
            cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
            DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
            Clone Distribution: Clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            The following repetitive elements were found in this cDNA
            sequence: 276-341, >GC rich#Low_complexity
            Seq primer: M13 FORWARD
            POLYA=Yes.
FEATURES             Location/Qualifiers
     1..633
     source            /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"

```

```

/clone="IMAGE:5839650"
/tissue_type="Chondrosarcoma"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP ED1"
/notes="Organ: Left Pubic Bone; Vector: p77T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP ED1 is a normalized cDNA library containing the following tissue(s): Chondrosarcoma cell line C55. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into p77T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is GCTCAAGGCT.
TAG_TISSUE=chondrosarcoma
TAG_LIB=UI-H-ED1
TAG_SEQ=CGTCAAGGCT"

```

## ORIGIN

```

Query Match      89.2%; Score 79.4; DB 3; Length 633;
Best Local Similarity 93.3%; Pred. No. 8.5e-08;
Matches 83; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CGGGAATCCCCCTTCAGTTCCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
   |||||
Db 91 CGGGAATCCCCCTTCAGTTCCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 32
   |||||

QY 61 AACAAACCCAAAAAAGAAAAAGAAAAA 89
   |||||
Db 31 AACAAACCGACTCACAAAAAAGAAAAA 3
   |||||

```

## RESULT 5

```

CB241372/c
LOCUS      CB241372      684 bp      mRNA      linear      EST 12-FEB-2003
DEFINITION UI-CF-FN0-afy-h-23-0-UI.s1 UI-CF-FN0 Homo sapiens cDNA clone
IMAGE:5839650 3', mRNA sequence.
ACCESSION  CB241372
VERSION    CB241372.1 GI:28363016
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homnidae; Homo.
REFERENCE  1 (bases 1 to 684)
            Bonaldo, M.F., Lennon, G. and Soares, M.B.
            Normalization and subtraction: two approaches to facilitate gene
            discovery
            Genome Res. 6 (9), 791-806 (1996)
            8889548
            Contact: McCray, PB
            McCray Lab
            University of Iowa
            2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
            Tel: 319 356 4866
            Fax: 319 356 7171
            Email: paul-mccray@uiowa.edu
            Tissue Procurement: Dr. M. J. Welsh, University of Iowa
            cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
            cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
            DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
            Clone Distribution: Researchers may obtain clones from Research
            Genetics (www.reagen.com) or from Open Biosystems
            (www.openbiosystems.com).
            The following repetitive elements were found in this cDNA
            sequence: 277-342, >GC rich#Low_complexity
            Seq primer: M13 FORWARD

```

```

POLYA=Yes.
FEATURES
  source      Location/Qualifiers
    1..684
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="UI-CF-FNO-afy-h-23-0-UI"
      /tissue_type="Human Lung Epithelial cells"
      /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
      /clone_lib="UI-CF-FNO"
      /notes="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
      modified polylinker; Site 1: Ecor I; Site 2: Not I;
      UI-CF-FNO is a subtracted cDNA library derived from two
      normalized Human lung epithelial cell libraries (EN1 and
      DU1) The library was subtracted according to according to
      Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
      1996. For additional information, contact:
      bento-soares@uiowa.edu
      TAG_TISSUE=Human Lung Epithelial Cell Lines untreated LPS
      6hr to LPS 24h
      TAG_LIB=UI-CF-FNO
      TAG_SEQ=CTGCTCAGGT"

ORIGIN
  Query Match      89.2%; Score 79.4; DB 6; Length 684;
  Best Local Similarity 93.3%; Pred. No. 8.5e-08;
  Matches 83; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCATGACTCGAATATCTGAAATGAAGAA 60
    |||||||
Db 92 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCATGACTCGAATATCTGAAATGAAGAA 33
    |||||||

QY 61 AACAAACCAAAAAAAAAAAAAAAAAAAAAA 89
    |||||||
Db 32 AACAAACCGACTCACAAAAAAAAAAAAAAAA 4

RESULT 6
CD364677/c
LOCUS
DEFINITION
  UI-H-FT2-bjm-i-14-0-UI.s1 NCI CGAP FT2 Homo sapiens cDNA clone
  UI-H-FT2-bjm-i-14-0-UI 3', mRNA sequence.
CD364677
VERSION
  CD364677.1 GI:31148767
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
  ORGANISM
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
    Hominidae; Homo.
REFERENCE
  1 (bases 1 to 691)
  NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
  Unpublished (1997)
  Contact: Robert Strausberg, Ph.D.
  Email: csapsb@mail.nih.gov
  Tissue Procurement: Dr. Gary W. Hunninghake, U of I
  cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
  cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
  DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
  Clone Distribution: Distribution information can be found at
  http://genome.uiowa.edu/distribution/cgap.html
  Seq primer: M13 FORWARD
POLYA=Yes.
FEATURES
  source      Location/Qualifiers
    1..691
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="UI-H-FT2-bjm-i-14-0-UI"
      /tissue_type="Alveolar Macrophage"
      /dev_stage="Adult"

```

```

/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP FT2"
/notes="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site 1: Ecor I; Site 2: Not I;
NCI CGAP FT2 is a subtracted cDNA library constructed from
a pool of 81 RNA samples from Alveolar Macrophages
challenged with different treatments. The mRNA samples
were a mixture of these conditions (times refer to
incubations following isolation by bronchoalveolar lavage)
(some normal donor macrophages were cultured in some of
the conditions, other donor macrophages in different
conditions). The mRNA samples were pooled for library
construction. Control 0 hours; control 3 hours; control 24
hours; LPS 100 ng/ml, 3 hours; LPS 100 ng/ml, 24 hours;
PMA 10 ng/ml, 3 hours; PMA 10 ng/ml, 24 hours; Klebsiella
moi 10, 3 hours; Klebsiella moi 10, 24 hours; Staph
aureus moi 10, 3 hours; Staph aureus moi 10, 24 hours;
Adenoviral vector (Ad5 CMV eGFP), moi 500, 3 hours;
Adenoviral vector (Ad5 CMV eGFP), moi 500, 24 hours; wt
adenovirus moi 500, 3 hours; wt adenovirus moi 500, 24
hours; Ad vector + LPS 3 hours; Ad vector + LPS 24 hours;
wt adenovirus + LPS 3 hours; wt adenovirus + LPS 24 hours.
The library was subtracted according to Bonaldo, Lennon
and Soares, Genome Research, 6:791-806, 1996. The tissue
was provided by Dr. Gary W. Hunninghake of the University
of Iowa.
TAG_TISSUE=Human Lung Alveolar Macrophage
TAG_LIB=UI-H-FT2
TAG_SEQ=GGCCATGCCG"

ORIGIN
  Query Match      89.2%; Score 79.4; DB 6; Length 691;
  Best Local Similarity 93.3%; Pred. No. 8.5e-08;
  Matches 83; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCATGACTCGAATATCTGAAATGAAGAA 60
    |||||||
Db 91 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCATGACTCGAATATCTGAAATGAAGAA 32
    |||||||

QY 61 AACAAACCAAAAAAAAAAAAAAAAAAAAAA 89
    |||||||
Db 31 AACAAACCGACTCACAAAAAAAAAAAAAAAA 3

RESULT 7
BM053069
LOCUS
DEFINITION
  1e69h03.v3 Melton Normalized Human Islet 4 N4-His 1 Homo sapiens
  cDNA clone IMAGE:5672308 5', mRNA sequence.
BM053069
VERSION
  BM053069.1 GI:16809019
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
  ORGANISM
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
    Hominidae; Homo.
REFERENCE
  1 (bases 1 to 313)
  Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,
  Lemishka, I., Scearce, M., Brestelli, J., Gradwohl, G., Clifton, S.,
  Hillier, L., Marra, M., Pape, D., Wyllie, T., Martin, J., Blissett, A.,
  Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J.,
  Cardenas, M., Gibbons, M., McCann, R., Cole, R., Tsagarisvilli, R.,
  Williams, T., Jackson, Y., and Bowers, Y.
  Endocrine Pancreas Consortium
  Unpublished (2000)
  Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
  Endocrine Pancreas Consortium
  Harvard University, Howard Hughes Medical Institute
  Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
  MA 02138
  Tel: 617-495-1812
  Fax: 617-495-8557

```

Email: dmelton@biohp.harvard.edu  
 Library was constructed by Dr. Douglas Melton DNA sequencing by:  
 Washington University Genome Sequencing Center For information on  
 obtaining a clone please contact: Juliana Brown  
 (brown@fas.harvard.edu) This sequence now available from the IMAGE  
 consortium, for clone orders contact: info@image.llnl.gov.  
 Location/Qualifiers  
 1. 313

## FEATURES

source

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5672308"
/sex="Both"
/tissue_type="Adult"
/dev_stage="Adult"
/lab_host="DH10B"
/clone_lib="Melton Normalized Human Islet 4 N4-HIS 1"
/note="Organ: Pancreas; Vector: pSPORT1; Site_1: Not 1; Site_2: Sal 1; Starting library constructed using SuperScript Plasmid Library kit (Life Technologies). cDNA made by oligo-dT priming. Size-selected by column fractionation; average insert size 1.08 kb. Library was amplified once on solid support and plasmid DNA from library was prepared. The library DNA was normalized by method #4 from Bonaldo, Lennon, and Soares 1996 Genome Research 6:791-806; 0.5 microgram single-stranded library Plasmid DNA was mixed with 5 micrograms PCR product representing library inserts and hybridized to an Ecot of 20. Single-stranded (unhybridized) plasmids were isolated by hydroxyapatite chromatography and used to make this library."

```

## ORIGIN

```

Query Match      84.7%; Score 75.4; DB 3; Length 313;
Best Local Similarity 92.9%; Pred.No. 6.8e-07;
Matches 79; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

```

Qy 1 CGGGAATCCCTTCAGTTCCTTTGAAAAAGTTCATGACTCGAATATCTGAAATGAAGAA 60
    |||||
Db 229 CGGGAATCCCTTCAGTTCCTTTGAAAAAGTTCATGACTCGAATATCTGAAATGAAGAA 288
    |||||

Qy 61 AACAAACCAAAAAAAGAA 85
    |||||
Db 289 AACAAACCGACTCACAAAAA 313
    |||||

```

```

RESULT 8
BF431778/c
LOCUS
DEFINITION
  nab50c04.x1 Soares NSF F8 9W OT PA_P_S1 Homo sapiens cDNA clone
  IMAGE:3269358 3', mRNA sequence.
ACCESSION
  BF431778
VERSION
  BF431778.1 GI:11443892
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
ORGANISM
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
  Homiidae; Homo.

```

```

REFERENCE
  1 (bases 1 to 292)
  NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS
  National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
TITLE
  Unpublished (1997)
JOURNAL
  Unpublished (1997)
COMMENT
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-remail.nih.gov
  This clone is available royalty-free through LLNL; contact the
  IMAGE Consortium (info@image.llnl.gov) for further information.
  Seq primer: -40UP from Gibco
  High quality sequence stop: 272.
  Location/Qualifiers
  1. 292
    /organism="Homo sapiens"

```

```

FEATURES
  source

```

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3269358"
/lab_host="DH10B"
/clone_lib="Soares NSF F8 9W OT PA_P_S1"
/note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not 1; Site_2: Eco RI; Equal amounts of plasmid DNA from five normalized libraries were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 5 libraries. The pools consisted of the following libraries and cloneIDs: Soares NDHSF pool 1: 309384-310919, 323208-325895 Soares NB2HP pool 1: 145032-147335, 147720-148103, 148872-149255, 15002 - 150407, 151176-152327 Soares NB2HP8-9W pool 1: 758280-760583, 772104-774407 Soares NBHPA pool 1: 304776-306311, 320136-322823, 326280-326663 Soares NBHOT pool 1: 723720-726407, 739080-740999 Subtraction by Bento Soares and M. Fatima Bonaldo."

```

## ORIGIN

```

Query Match      82.0%; Score 73; DB 2; Length 292;
Best Local Similarity 93.8%; Pred.No. 2.3e-06;
Matches 76; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

## RESULT 9

AA903087/c

LOCUS

DEFINITION

AA903087

AA903087.1

GI:3038210

EST.

KEYWORDS

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 310)

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

unknown library type

Seq primer: -40ml3 fwd. ET from Amersham

High quality sequence stop: 295.

Location/Qualifiers

1. 310

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:1516999"

/tissue\_type="leiomyosarcoma"

/lab\_host="DH10B"

/clone\_lib="NCI CGAP Lei2"

/note="Organ: soft tissue; Vector: pT7T3D-Pac (Pharmacia)

with a modified polylinker; Site\_1: Not 1; Site\_2: Eco RI;

1st strand cDNA was primed with a Not 1 - oligo(dT) primer

15' -AAGTGAAGATTCGCGCGCAATCGTTTTTTTTTTTTTT-3',

double-stranded cDNA was ligated to Eco RI adaptors



Best Local Similarity 93.8%; Pred. No. 2.2e-06;  
Matches 76; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60  
|||||  
Db 81 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 22  
|||||  
Qy 61 AACAAACCAAAAAA 81  
|||||  
Db 21 AACAAACCGACTCACAAAAA 1  
|||||

RESULT 12  
BF002331/c  
LOCUS  
DEFINITION 716 bp mRNA linear EST 06-OCT-2000  
7h22b08.x1 NCI CGAP Col6 Homo sapiens cDNA clone IMAGE:3316695 3'  
similar to TR:O00465 O00466 K12 PROTEIN PRECURSOR. ;, mRNA  
sequence.

ACCESSION BF002331 GI:10702606  
VERSION BF002331  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
1 (bases 1 to 716)

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-re@mail.nih.gov  
Tissue Procurement: Ilan Kirsch, M.D., Michael R. Emmert-Buck,  
M.D., Ph.D.  
cDNA Library Preparation: M. Bento Soares, Ph.D.  
DNA Sequencing by: Greg Lennon, Ph.D.  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL, send email to:  
[info@image.llnl.gov](mailto:info@image.llnl.gov)

Seq primer: -40UP from Gibco  
High quality sequence stop: 462.

FEATURES Location/Qualifiers

1..716

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3316695"  
/tissue\_type="colon tumor, RER"  
/lab\_host="DH10B"  
/clone\_lib="NCI CGAP Col6"

/note="Organ: colon; Vector: p7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not 1; Site 2: Eco RI; Plasmid DNA from the normalized library NCI CGAP Col6 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 1057416-1061255, and 1144584-1145351). Subtraction by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 80.9%; Score 72; DB 2; Length 716;  
Best Local Similarity 93.8%; Pred. No. 3.6e-06;  
Matches 75; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60  
|||||  
Db 80 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 21  
|||||  
Qy 61 AACAAACCAAAAAA 80  
|||||

Db 20 AACAAACCGACTCACAAAAA 1  
|||||

RESULT 13

AI480347/c

LOCUS

DEFINITION 445 bp mRNA linear EST 14-APR-1999

AI480347

ACCESSION

VERSION AI480347

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 445)

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-re@mail.nih.gov

Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.

Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D.

DNA Sequencing by: Greg Lennon, Ph.D.

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

[www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)

Insert Length: 533 Std Error: 0.00

Seq primer: -40UP from Gibco.

FEATURES Location/Qualifiers

1..445

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clones="IMAGE:2161648"

/lab\_host="DH10B"

/clone\_lib="NCI CGAP Kid11"

/note="Organ: kidney; Vector: p7T3D-Pac (Pharmacia) with

a modified polylinker; Site 1: Not 1; Site 2: Eco RI;

Plasmid DNA from the normalized library NCI CGAP Kid3 was

prepared, and ss circles were made in vitro. Following HAP

purification, this DNA was used as tracer in a subtractive

hybridization reaction. The driver was PCR-amplified cDNAs

from a pool of 5,000 clones made from the same library

(cloneIDs 1322376-1323911, 1456007-1456775, and

1500552-1502855). Subtraction by Bento Soares and M.

Fatima Bonaldo."

ORIGIN

Query Match 80.4%; Score 71.6; DB 1; Length 445;

Best Local Similarity 94.9%; Pred. No. 4.6e-06;

Matches 74; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60

|||||

Db 78 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 19

|||||

Qy 61 AACAAACCAAAAAA 78

|||||

Db 18 AACAAACCAAAAAA 1

|||||

RESULT 14

BU689727/c

LOCUS

DEFINITION 687 bp mRNA linear EST 07-OCT-2002

BU689727

ACCESSION

VERSION BU689727

UI-CF-FNO-aet-k-09-0-UI.81

UI-CF-FNO-aet-k-09-0-UI 3', mRNA sequence.

BU689727

BU689727.1 GI:23547747



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 11:34:11 ; Search time 60.7533 Seconds  
(without alignments)  
455.919 Million cell updates/sec

Title: US-09-980-046B-4  
Perfect score: 89  
Sequence: 1 cgggaatcccccttcagtc.....aaaaaaaaaagaaaaa 89

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3289935 seqs, 155610033 residues

Total number of hits satisfying chosen parameters: 6579870

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA New:\*  
1: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq3:\*  
10: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	41.4	46.5	2066	US-10-523-588-2	Sequence 2, Appli
2	36.8	41.3	2027	US-10-131-826A-175	Sequence 175, App
3	36.2	40.7	2668	US-10-131-826A-511	Sequence 511, App
4	36	40.4	1941	US-10-131-826A-165	Sequence 165, App
5	36	40.4	2625	US-11-102-978-12	Sequence 12, Appl
6	35.8	40.2	1872	US-10-750-185-31235	Sequence 31235, A
7	35.2	39.6	171427	US-11-112-908-60	Sequence 60, Appl
8	35	39.3	644	US-11-102-240-55	Sequence 55, Appl
9	34.6	38.9	778	US-10-986-501-38	Sequence 38, Appl
10	34.6	38.9	5503	US-11-112-908-6	Sequence 6, Appli
11	34.4	38.7	1968	US-10-131-826A-163	Sequence 163, App
12	34.4	38.7	2537	US-11-102-978-8	Sequence 8, Appli
13	34.4	38.7	3124	US-10-750-185-38847	Sequence 38847, A
14	34.2	38.4	1478	US-10-909-125-1744	Sequence 1744, Ap
15	34.2	38.4	4740	US-11-080-026-3	Sequence 3, Appli
16	34	38.2	485	US-11-102-240-43	Sequence 43, Appl
17	33.8	38.0	687	US-10-986-501-107	Sequence 107, App
18	33.8	38.0	2091	US-10-276-233A-19	Sequence 19, Appl
19	33.8	38.0	2333	US-10-131-826A-345	Sequence 345, App
20	33.8	38.0	3501	US-10-131-826A-37	Sequence 37, Appl
21	33.6	37.8	783	US-10-131-826A-321	Sequence 321, App
22	33.6	37.8	1979	US-10-967-527A-31	Sequence 31, Appl
23	33.4	37.5	2037	US-10-909-125-799	Sequence 799, App

24	33.4	37.5	2207	6	US-10-131-826A-137	Sequence 137, App
25	33.2	37.3	2120	6	US-10-689-742-163	Sequence 163, App
26	33.2	37.3	2773	6	US-10-689-742-43	Sequence 43, Appl
27	33	37.1	1073	7	US-11-102-240-95	Sequence 95, Appl
28	33	37.1	1108	6	US-10-689-742-209	Sequence 209, App
29	33	37.1	1234	7	US-11-102-240-63	Sequence 63, Appl
30	33	37.1	2822	6	US-10-131-826A-305	Sequence 305, App
31	33	37.1	6410	7	US-11-108-172-1094	Sequence 1094, Ap
32	32.8	36.9	2487	6	US-10-689-742-165	Sequence 165, App
33	32.8	36.9	2570	6	US-10-131-826A-539	Sequence 539, App
34	32.8	36.9	2570	7	US-11-102-240-167	Sequence 167, App
35	32.8	36.9	12482	9	US-11-090-878-25	Sequence 25, Appl
36	32.6	36.6	755	6	US-10-131-826A-153	Sequence 153, App
37	32.6	36.6	1240	6	US-10-689-742-81	Sequence 81, Appl
38	32.6	36.6	2436	7	US-11-102-240-99	Sequence 99, Appl
39	32.6	36.6	2463	7	US-11-108-528-3	Sequence 3, Appli
40	32.6	36.6	3604	6	US-10-750-185-53243	Sequence 53243, A
41	32.6	36.6	4606	6	US-10-750-185-50060	Sequence 50060, A
42	32.4	36.4	286	6	US-10-986-501-95	Sequence 95, Appl
43	32.4	36.4	1573	6	US-10-689-742-187	Sequence 187, App
44	32.4	36.4	1844	6	US-10-131-826A-475	Sequence 475, App
45	32.4	36.4	2476	6	US-10-131-826A-489	Sequence 489, App

ALIGNMENTS

RESULT 1  
US-10-523-588-2  
; Sequence 2, Application US/10523588  
; Publication No. US20050251870A1  
; GENERAL INFORMATION:  
; APPLICANT: EXELIXIS, INC.  
; TITLE OF INVENTION: CSNK1GS AS MODIFIERS OF THE P21 PATHWAY AND METHODS OF USE  
; FILE REFERENCE: EX03-057C-US  
; CURRENT APPLICATION NUMBER: US/10/523,588  
; CURRENT FILING DATE: 2005-02-04  
; PRIOR APPLICATION NUMBER: US 60/401,739  
; PRIOR FILING DATE: 2002-08-07  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 2  
; LENGTH: 2066  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; US-10-523-588-2

Query Match	46.5%;	Score 41.4;	DB 6;	Length 2066;
Best Local Similarity	68.7%;	Pred. No. 0.13;		
Mismatches	57;	Conservative	0;	Mismatches 26; Indels 0; Gaps 0;
Qy	7	TCCCCCTTCAGTCTCTTTGAAAAAGCTTCATGACTCGAATATCTGAAATGAGAAAAACAAA	66	
Db	1953	TGCACCTTTTGTAGTTTTTTTAAAAAACAACAAAACATGGCGATGCGAAAAA	2012	
Qy	67	CCAAAAAAGAAAAAGAAAAA	89	
Db	2013	AAAAAAGAAAAAAGAAAAA	2035	

RESULT 2  
US-10-131-826A-175  
; Sequence 175, Application US/10131826A  
; Publication No. US20050245730A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Deanoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey





```
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C128
; CURRENT APPLICATION NUMBER: US/10/131,826A
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; Remaining Prior Application data removed - See File wrapper or PALM.
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 165
; LENGTH: 1941
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-131-826A-165

Query Match 40.4%; Score 36; DB 6; Length 1941;
Best Local Similarity 67.1%; Pred. No. 2.1;
Matches 51; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

Qy 14 TCAGTCTCTTGAAGAAAGTTCCATCGAATATCTGAATCTGAATCAAGAAACCAACCAAAAA 73
Db 1860 TCTGTGACTGATTAAGTTCCATTTTGTGGAAAAAAGAAAAAAGAAAAAAGAAAAA 1919

Qy 74 AAAAAAAGAAAAA 89
Db 1920 AAAAAAAGAAAAA 1935

RESULT 5
US-11-102-978-12
; Sequence 12, Application US/11102978
; Publication No. US20050250142A1
; GENERAL INFORMATION:
; APPLICANT: University of Utah Technology Transfer Office
; APPLICANT: University of Utah Research Foundation
; TITLE OF INVENTION: Diagnosis and Treatment of Herpes Simplex Virus Disease
; FILE REFERENCE: 0274-5537.1US
; CURRENT APPLICATION NUMBER: US/11/102,978
; CURRENT FILING DATE: 2005-04-11
; PRIOR APPLICATION NUMBER: PC/US2003/033152
; PRIOR FILING DATE: 2003-10-18
; PRIOR APPLICATION NUMBER: 60/419,576
; PRIOR FILING DATE: 2002-10-18
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12
; LENGTH: 2625
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: gene
; LOCATION: (1)..(2625)
; OTHER INFORMATION: CHODL
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (392)..(1213)
```

```
; OTHER INFORMATION: The gene is CHODL, encoding chondrolectin, NM024944.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (506)..(928)
; OTHER INFORMATION: C-type lectin (CTL) or carbohydrate-recognition domain (CRD)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (539)..(931)
; OTHER INFORMATION: C-type lectin domain (CRD), including both the long and short
; FEATURE: form C-type.
; NAME/KEY: allele
; LOCATION: (1975)..(1975)
; OTHER INFORMATION: Comprises A or G.
; FEATURE:
; NAME/KEY: allele
; LOCATION: (2032)..(2032)
; OTHER INFORMATION: Comprises A or G.
; FEATURE:
; NAME/KEY: allele
; LOCATION: (2121)..(2121)
; OTHER INFORMATION: Comprises A or T.
; FEATURE:
; NAME/KEY: allele
; LOCATION: (2349)..(2349)
; OTHER INFORMATION: Comprises G or T.
; FEATURE:
; NAME/KEY: allele
; LOCATION: (2403)..(2404)
; OTHER INFORMATION: TCTA
; FEATURE:
; NAME/KEY: allele
; LOCATION: (2543)..(2543)
; OTHER INFORMATION: Comprises A or G.
; US-11-102-978-12

Query Match 40.4%; Score 36; DB 7; Length 2625;
Best Local Similarity 64.3%; Pred. No. 2.1;
Matches 54; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

Qy 6 ATCCCTTCAGTCTCTTGAAGAAAGTTCCATCGAATATCTGAATCAAGAAACCAACCA 65
Db 2496 ATGTCCTGTGCTCTCTTTTAAACCAATAAAGAGTTCTTGTTCGAAAGAAAAA 2555

Qy 66 ACCAAAAAAGAAAAA 89
Db 2556 AAAAAAAGAAAAA 2579

RESULT 6
US-10-750-185-31235
; Sequence 31235, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DeNISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 31235
; LENGTH: 1872
; TYPE: DNA
; ORGANISM: Bovine 19866881231610
```

US-10-750-185-31235

Query Match 40.2%; Score 35.8; DB 6; Length 1872;  
Best Local Similarity 63.2%; Pred. No. 2.3;  
Matches 55; Conservative 0; Mismatches 32; Indels 0; Gaps 0;  
QY 3 GGAATCCCTTCAGTTCCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAAAA 62  
DB 80 GAAAGACAGTAATAATTTTAAAAAAGATCAACCAAGAAAAATACAAATTTGAAACACA 139  
QY 63 CAAACCAAAAAAAGAAAAAAGAAAAA 89  
DB 140 AAGAGAAAAAAGAAAAAGAGAAAAA 166

RESULT 7

US-11-112-908-60/c  
; Sequence 60, Application US/11112908  
; Publication No. US20050260659A1  
; GENERAL INFORMATION:  
; APPLICANT: Harris, Cole  
; TITLE OF INVENTION: Breast Cancer Biomarkers  
; FILE REFERENCE: 04-164-US  
; CURRENT APPLICATION NUMBER: US/11/112,908  
; CURRENT FILING DATE: 2005-04-22  
; PRIOR APPLICATION NUMBER: US 60/564,758  
; PRIOR FILING DATE: 2004-04-23  
; PRIOR APPLICATION NUMBER: US 60/575,978  
; PRIOR FILING DATE: 2004-06-01  
; PRIOR APPLICATION NUMBER: US 60/631,702  
; PRIOR FILING DATE: 2004-11-30  
; PRIOR APPLICATION NUMBER: US 60/633,826  
; PRIOR FILING DATE: 2004-12-07  
; NUMBER OF SEQ ID NOS: 511  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 60  
; LENGTH: 171427  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-11-112-908-60

Query Match 39.6%; Score 35.2; DB 7; Length 171427;  
Best Local Similarity 76.8%; Pred. No. 4.5;  
Matches 43; Conservative 0; Mismatches 13; Indels 0; Gaps 0;  
QY 34 CATGACTCGAATATCTGAAATGAAGAAAAACCAACCAAAAAAAGAAAAA 89  
DB 104488 CAAGAATCCCATTTTCAGAAAAAAGAAAAAAGAAAAAAGAAAAA 104433

RESULT 8

US-11-102-240-55  
; Sequence 55, Application US/11102240  
; Publication No. US20050260647A1  
; GENERAL INFORMATION:  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, Christopher J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Wood, William I.  
; TITLE OF INVENTION: ANTIBODIES TO POLYPEPTIDES ENCODED BY A NUCLEIC ACID UNDEREXPRES  
; FILE REFERENCE: P3230R1C106C  
; CURRENT APPLICATION NUMBER: US/11/102,240  
; CURRENT FILING DATE: 2005-04-08  
; PRIOR APPLICATION NUMBER: 10/063662  
; PRIOR FILING DATE: 2002-05-07  
; PRIOR APPLICATION NUMBER: 10/006867  
; PRIOR FILING DATE: 2001-12-06  
; PRIOR APPLICATION NUMBER: PCT/US00/23328  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: 60/170262

; PRIOR FILING DATE: 199-12-09  
; NUMBER OF SEQ ID NOS: 170  
; SEQ ID NO 55  
; LENGTH: 644  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-11-102-240-55

Query Match 39.2%; Score 35; DB 7; Length 644;  
Best Local Similarity 63.9%; Pred. No. 3.2;  
Matches 53; Conservative 0; Mismatches 30; Indels 0; Gaps 0;  
QY 7 TCCCTTCAGTTCCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAAAAA 66  
DB 532 TCCACTGAATTCCTCATGAATACAACTATTTCAGCAACAGCAAAAAA 591  
QY 67 CCAAAAAAAGAAAAAAGAAAAA 89  
DB 592 AAAAAAAGAAAAAAGAAAAA 614

RESULT 9

US-10-986-501-38  
; Sequence 38, Application US/10986501  
; Publication No. US20050244845A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: 90 Human Secreted Proteins  
; FILE REFERENCE: P2013P2C1  
; CURRENT APPLICATION NUMBER: US/10/986,501  
; CURRENT FILING DATE: 2004-11-12  
; PRIOR APPLICATION NUMBER: US/10/621,363  
; PRIOR FILING DATE: 2003-07-18  
; PRIOR APPLICATION NUMBER: 09/969,730  
; PRIOR FILING DATE: 2001-10-06  
; PRIOR APPLICATION NUMBER: 09/774,639  
; PRIOR FILING DATE: 2001-02-01  
; PRIOR APPLICATION NUMBER: 60/238,291  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 09/244,112  
; PRIOR FILING DATE: 1999-02-04  
; PRIOR APPLICATION NUMBER: PCT/US98/16235  
; PRIOR FILING DATE: 1998-08-04  
; PRIOR APPLICATION NUMBER: 60/056,371  
; PRIOR FILING DATE: 1997-08-19  
; PRIOR APPLICATION NUMBER: 60/056,732  
; PRIOR FILING DATE: 1997-08-19  
; PRIOR APPLICATION NUMBER: 60/056,366  
; PRIOR FILING DATE: 1997-08-19  
; PRIOR APPLICATION NUMBER: 60/056,364  
; PRIOR FILING DATE: 1997-08-19  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 373  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 38  
; LENGTH: 778  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-986-501-38

Query Match 38.9%; Score 34.6; DB 6; Length 778;  
Best Local Similarity 67.1%; Pred. No. 3.9;  
Matches 49; Conservative 0; Mismatches 24; Indels 0; Gaps 0;  
QY 17 GTTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAAAAAAGAAAAA 76  
DB 695 GTCATTGTAATAAAGCCTCTGCTCTCAGATGTATAAAAAAAGAAAAAAGAAAAA 754  
QY 77 AAAAAAAGAAAAA 89  
DB 755 AAAAAAAGAAAAA 767

```
RESULT 10
US-11-112-908-6
; Sequence 6, Application US/11112908
; Publication No. US20050260659A1
; GENERAL INFORMATION:
; APPLICANT: Harris, Cole
; APPLICANT: Davis, Lisa M.
; TITLE OF INVENTION: Breast Cancer Biomarkers
; FILE REFERENCE: 04-164-US
; CURRENT APPLICATION NUMBER: US/11/112,908
; CURRENT FILING DATE: 2005-04-22
; PRIOR APPLICATION NUMBER: US 60/564,758
; PRIOR FILING DATE: 2004-04-23
; PRIOR APPLICATION NUMBER: US 60/575,978
; PRIOR FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US 60/631,702
; PRIOR FILING DATE: 2004-11-30
; PRIOR APPLICATION NUMBER: US 60/633,826
; PRIOR FILING DATE: 2004-12-07
; NUMBER OF SEQ ID NOS: 511
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 6
; LENGTH: 5503
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-112-908-6

Query Match      38.9%; Score 34.6; DB 7; Length 5503;
Best Local Similarity 67.1%; Pred. No. 4.6; Mismatches 24; Indels 0; Gaps 0;
Matches 49; Conservative 0;

Qy 17 GTTCTTTGAAAAAGTTCCATGACTCGAATATCTCGAATGGAAGAAAAACAAACAAAAA 76
Db 5415 GGTTTTAAAGAACCTCAGAAGCTATTCTTAAAGAAAAAAGAAAAAAGAAAAA 5474

Qy 77 AAAAAAGAAAAA 89
Db 5475 AAAAAAGAAAAA 5487

RESULT 11
US-10-131-826A-163
; Sequence 163, Application US/10131826A
; Publication No. US20050245730A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330RIC128
; CURRENT APPLICATION NUMBER: US/10/131,826A
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
```

```
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 163
; LENGTH: 1968
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-131-826A-163

Query Match      38.7%; Score 34.4; DB 6; Length 1968;
Best Local Similarity 63.1%; Pred. No. 4.7; Mismatches 31; Indels 0; Gaps 0;
Matches 53; Conservative 0;

Qy 6 ATCCCCCTTCAGTTCTTTGAAAAAGTTTCCATGACTCGAATATCTCGAATGGAAGAAAAA 65
Db 1883 ATCTATTTCATCCTCGAATAATAAGAAATATATTTCACCTCTAAAAAAGAAAAA 1942

Qy 66 ACCAAAAAAGAAAAAAGAAAAA 89
Db 1943 AAAAAAAGAAAAAAGAAAAA 1966

RESULT 12
US-11-102-978-8
; Sequence 8, Application US/11102978
; Publication No. US20050250142A1
; GENERAL INFORMATION:
; APPLICANT: University of Utah Technology Transfer Office
; APPLICANT: University of Utah Research Foundation
; TITLE OF INVENTION: Diagnosis and Treatment of Herpes Simplex Virus Disease
; FILE REFERENCE: 0274-5537.1US
; CURRENT APPLICATION NUMBER: US/11/102,978
; CURRENT FILING DATE: 2005-04-11
; PRIOR APPLICATION NUMBER: PCT/US2003/033152
; PRIOR FILING DATE: 2003-10-18
; PRIOR APPLICATION NUMBER: 60/419,576
; PRIOR FILING DATE: 2002-10-18
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 8
; LENGTH: 2537
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (102)..(1199)
; OTHER INFORMATION: Homo sapiens coxsackie virus and adenovirus receptor (CXADR),
; OTHER INFORMATION: mRNA, genbank NM001338.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (186)..(515)
; OTHER INFORMATION: IG, Immunoglobulin
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (207)..(467)
; OTHER INFORMATION: IGV, Immunoglobulin V-Type
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (558)..(758)
; OTHER INFORMATION: Igc2, Immunoglobulin C-2 Type
```

/ FEATURE:  
/ NAME/KEY: misc feature  
/ LOCATION: (564)..(788)  
/ OTHER INFORMATION: IG, Immunoglobulin  
/ FEATURE:  
/ NAME/KEY: misc feature  
/ LOCATION: (750)..(750)  
/ OTHER INFORMATION: includes an A or G nucleotide  
/ FEATURE:  
/ NAME/KEY: allele  
/ LOCATION: (969)..(969)  
/ OTHER INFORMATION: may comprise G or A.  
/ FEATURE:  
/ NAME/KEY: allele  
/ LOCATION: (1148)..(1148)  
/ OTHER INFORMATION: may comprise G or A.  
/ FEATURE:  
/ NAME/KEY: Unsure  
/ LOCATION: (1148)..(1148)  
/ OTHER INFORMATION: Comprises A or C.  
/ FEATURE:  
/ NAME/KEY: Unsure  
/ LOCATION: (1210)..(1210)  
/ OTHER INFORMATION: Comprises T or C.  
/ FEATURE:  
/ NAME/KEY: allele  
/ LOCATION: (1269)..(1269)  
/ OTHER INFORMATION: may comprise G or T.  
/ FEATURE:  
/ NAME/KEY: allele  
/ LOCATION: (1278)..(1278)  
/ OTHER INFORMATION: may comprise G or T.  
/ FEATURE:  
/ NAME/KEY: allele  
/ LOCATION: (1293)..(1293)  
/ OTHER INFORMATION: may comprise G or A.  
/ FEATURE:  
/ NAME/KEY: allele  
/ LOCATION: (1297)..(1297)  
/ OTHER INFORMATION: may comprise G or A.  
/ FEATURE:  
/ NAME/KEY: allele  
/ LOCATION: (1298)..(1298)  
/ OTHER INFORMATION: may comprise C or T.  
/ FEATURE:  
/ NAME/KEY: allele  
/ LOCATION: (1322)..(1322)  
/ OTHER INFORMATION: may comprise A or G.  
/ FEATURE:  
/ NAME/KEY: allele  
/ LOCATION: (1365)..(1365)  
/ OTHER INFORMATION: may comprise A or G.  
/ FEATURE:  
/ NAME/KEY: allele  
/ LOCATION: (1376)..(1376)  
/ OTHER INFORMATION: may comprise A or G.  
/ FEATURE:  
/ NAME/KEY: allele  
/ LOCATION: (1381)..(1381)  
/ OTHER INFORMATION: may comprise C or T.  
/ FEATURE:  
/ NAME/KEY: allele  
/ LOCATION: (1397)..(1397)  
/ OTHER INFORMATION: may comprise C or A.  
/ FEATURE:  
/ NAME/KEY: allele  
/ LOCATION: (1449)..(1449)  
/ OTHER INFORMATION: may comprise G or T.  
/ FEATURE:  
/ NAME/KEY: allele  
/ LOCATION: (1827)..(1827)  
/ OTHER INFORMATION: may comprise G or A.  
/ FEATURE:

/ NAME/KEY: allele  
/ LOCATION: (1866)..(1866)  
/ OTHER INFORMATION: may comprise C or T.  
/ FEATURE:  
/ NAME/KEY: Unsure  
/ LOCATION: (1976)..(1976)  
/ OTHER INFORMATION: Comprises A or G.  
/ FEATURE:  
/ NAME/KEY: allele  
/ LOCATION: (2007)..(2007)  
/ OTHER INFORMATION: may comprise C or T.  
/ FEATURE:  
/ NAME/KEY: allele  
/ LOCATION: (2095)..(2095)  
/ OTHER INFORMATION: may comprise A or G.  
/ FEATURE:  
/ NAME/KEY: allele  
/ LOCATION: (2106)..(2106)  
/ OTHER INFORMATION: may comprise A or C.  
/ FEATURE:  
/ NAME/KEY: allele  
/ LOCATION: (2154)..(2154)  
/ OTHER INFORMATION: may comprise T or C.  
/ FEATURE:  
/ NAME/KEY: allele  
/ LOCATION: (2159)..(2159)  
/ OTHER INFORMATION: may comprise T or C.  
/ FEATURE:  
/ NAME/KEY: allele  
/ LOCATION: (2192)..(2192)  
/ OTHER INFORMATION: may comprise T or C.  
/ FEATURE:  
/ NAME/KEY: allele  
/ LOCATION: (2193)..(2193)  
/ OTHER INFORMATION: may comprise A or G.  
/ FEATURE:  
/ NAME/KEY: allele  
/ LOCATION: (2271)..(2271)  
/ OTHER INFORMATION: may comprise A or C.  
/ FEATURE:  
/ NAME/KEY: Unsure  
/ LOCATION: (2294)..(2294)  
/ OTHER INFORMATION: Comprises T or C.  
/ FEATURE:  
/ NAME/KEY: Unsure  
/ LOCATION: (2328)..(2328)  
/ OTHER INFORMATION: Comprises T or C.  
/ FEATURE:  
/ NAME/KEY: Unsure  
/ LOCATION: (2340)..(2340)  
/ OTHER INFORMATION: Comprises T or A.  
/ FEATURE:  
/ NAME/KEY: Unsure  
/ LOCATION: (2346)..(2346)  
/ OTHER INFORMATION: Comprises G or A.  
/ FEATURE:  
/ NAME/KEY: Unsure  
/ LOCATION: (2354)..(2354)  
/ OTHER INFORMATION: Comprises G or A.  
/ FEATURE:  
/ NAME/KEY: Unsure  
/ LOCATION: (2374)..(2374)  
/ OTHER INFORMATION: Comprises G or A.  
/ FEATURE:  
/ NAME/KEY: Unsure  
/ LOCATION: (2413)..(2413)  
/ OTHER INFORMATION: Comprises C or T.  
/ FEATURE:  
/ NAME/KEY: allele  
/ LOCATION: (2440)..(2440)  
/ OTHER INFORMATION: may comprise A or G.  
/ FEATURE:

US-11-102-978-8

Query Match 38.7%; Score 34.4; DB 7; Length 2537;  
Best Local Similarity 65.8%; Pred. No. 4.8;  
Matches 50; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

Qy 14 TCAGTCTCTTGAAGAGTCCATGACTCGATATCTGAATGGAAGAAACCAACCAAAA 73  
Db 2427 TTAGTTTGTGAAGTGATTTATCTTAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 2486

Qy 74 AAAAAAAGAAAAA 89  
Db 2487 AAAAAAAGAAAAA 2502

## RESULT 13

US-10-750-185-38847  
; Sequence 38847, Application US/10750185  
; Publication No. US20050260603A1

## GENERAL INFORMATION:

; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MM1100-2  
; CURRENT APPLICATION NUMBER: US/10/750,185  
; CURRENT FILING DATE: 2003-12-31  
; PRIOR APPLICATION NUMBER: US 60/437,482  
; PRIOR FILING DATE: 2002-12-31  
; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 38847  
; LENGTH: 3124  
; TYPE: DNA  
; ORGANISM: Bovine 19866808035157

US-10-750-185-38847

Query Match 38.7%; Score 34.4; DB 6; Length 3124;  
Best Local Similarity 63.1%; Pred. No. 4.9;  
Matches 53; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

Qy 5 AATCCCCCTCAGTCTTGTGAAAAAGTCCATGACTCGAATATCTGAATGAAGAAAA 64  
Db 2640 AATACCTGAAAAATATTTTAAAAATTTTCAACACGACTTGCTATATAAGAAATC 2699

Qy 65 AACCAAAAAAAGAAAAA 88  
Db 2700 ATCCAAAAAAGAAAAA 2723

## RESULT 14

US-10-909-125-1744  
; Sequence 1744, Application US/10909125  
; Publication No. US20050261218A1

## GENERAL INFORMATION:

; APPLICANT: Esau, Christine  
; APPLICANT: Lollo, Bridget  
; APPLICANT: Bennett, C. Frank  
; APPLICANT: Freier, Susan M.  
; APPLICANT: Griffey, Richard H.  
; APPLICANT: Baker, Brenda F.  
; APPLICANT: Vickers, Timothy  
; APPLICANT: Marcusson, Eric G.  
; APPLICANT: Koller, Erich  
; APPLICANT: Swayze, Eric  
; APPLICANT: Jain, Ravi  
; APPLICANT: Bhat, Balkrishen  
; APPLICANT: Feralta, Eigen  
; TITLE OF INVENTION: Oligomeric Compounds And Compositions For Use In Modulation  
; FILE REFERENCE: ISIS0080-100 (CORE0016US)

; CURRENT APPLICATION NUMBER: US/10/909,125  
; CURRENT FILING DATE: 2004-07-30  
; PRIOR APPLICATION NUMBER: US 60/492,056  
; PRIOR FILING DATE: 2003-07-31

; PRIOR APPLICATION NUMBER: US 60/516,303  
; PRIOR FILING DATE: 2003-10-31  
; PRIOR APPLICATION NUMBER: US 60/531,596  
; PRIOR FILING DATE: 2003-12-19  
; PRIOR APPLICATION NUMBER: US 60/562,417  
; PRIOR FILING DATE: 2004-04-14

; NUMBER OF SEQ ID NOS: 2184  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1744  
; LENGTH: 1478  
; TYPE: DNA  
; ORGANISM: H. sapiens  
US-10-909-125-1744

Query Match 38.4%; Score 34.2; DB 6; Length 1478;  
Best Local Similarity 64.6%; Pred. No. 5.1;  
Matches 51; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 10 CCCTTCAGTCTTGTGAAAAAGTCCATGACTCGAATATCTGAATGAAGAAACCAACCA 69  
Db 1400 CCTTCTGATTAAGAAATTTTATTACCTAAAGAAAAAAGAAAAAAGAAAAA 1459

Qy 70 AAAAAAAGAAAAA 88  
Db 1460 AAAAAAAGAAAAA 1478

## RESULT 15

US-11-080-026-3  
; Sequence 3, Application US/11080026  
; Publication No. US20050260192A1

## GENERAL INFORMATION:

; APPLICANT: Springer, Timothy A.  
; APPLICANT: Shimaoka, Motomu  
; APPLICANT: Lu, Chafen  
; TITLE OF INVENTION: MODIFIED POLYPEPTIDES STABILIZED IN A  
; FILE REFERENCE: CFBF-P02-021  
; CURRENT APPLICATION NUMBER: US/11/080,026  
; CURRENT FILING DATE: 2005-03-15  
; PRIOR APPLICATION NUMBER: 09/945,265  
; PRIOR FILING DATE: 2001-08-31  
; PRIOR APPLICATION NUMBER: US 60/229,700  
; PRIOR FILING DATE: 2000-09-01  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 4740  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-11-080-026-3

Query Match 38.4%; Score 34.2; DB 7; Length 4740;  
Best Local Similarity 76.4%; Pred. No. 5.6;  
Matches 42; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 35 ATGACTCGAATATCTGAATGAAGAAACCAACCAAGAAAAAAGAAAAA 89  
Db 4682 ATAAATCAATATATGTCAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 4736

Search completed: December 6, 2005, 23:16:51  
Job time : 61.7533 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 21:43:03 ; Search time 778.261 Seconds  
(without alignments)  
6427.437 Million cell updates/sec

Title: US-09-980-046B-5  
Perfect score: 88  
Sequence: 1 cggagtcgagtcctgtatca.....cgtgttcttggtagacaaaa 88

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

#### Database :

GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_in:\*  
3: gb\_env:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pt:\*  
9: gb\_ro:\*  
10: gb\_scs:\*  
11: gb\_by:\*  
12: gb\_un:\*  
13: gb\_vl:\*  
14: gb\_brg:\*  
15: gb\_pl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	81.6	92.7	645	BC022439	BC022439 Homo sapi
2	80	90.9	251	BV167599	BV167599 sqm5991
3	80	90.9	467	CQ415312	CQ415312 Sequence
4	80	90.9	536	CQ433071	CQ433071 Sequence
5	80	90.9	541	CQ425463	CQ425463 Sequence
6	80	90.9	630	BC070243	BC070243 Homo sapi
7	80	90.9	642	BD237072	BD237072 Compounds
8	80	90.9	642	AR225472	AR225472 Sequence
9	80	90.9	642	AR562895	AR562895 Sequence
10	80	90.9	642	AX321588	AX321588 Sequence
11	80	90.9	222085	AC144988	AC144988 Gorilla g
12	78.4	89.1	637	BC006794	BC006794 Homo sapi
13	78.4	89.1	695	AX281855	AX281855 Sequence
14	78.4	89.1	749	BC008417	BC008417 Homo sapi
15	76.8	87.3	451	BD222194	BD222194 Human nuc
16	76.8	87.3	451	AX014315	AX014315 Sequence
17	75.8	86.1	118	CQ670675	CQ670675 Sequence
18	75.8	86.1	594	CQ923967	CQ923967 Sequence

c	19	75.8	86.1	668	6	CQ434132	CQ434132 Sequence
	20	75.8	86.1	808	6	BD190297	BD190297 Elki phos
	21	75.8	86.1	808	6	CQ812348	CQ812348 Sequence
	22	75.8	86.1	808	6	CS036376	CS036376 Sequence
	23	75.8	86.1	808	6	CS045328	CS045328 Sequence
	24	75.8	86.1	808	6	CS073286	CS073286 Sequence
	25	75.8	86.1	808	6	CS108221	CS108221 Sequence
	26	75.8	86.1	808	6	AR380734	AR380734 Sequence
	27	75.8	86.1	808	6	AR489210	AR489210 Sequence
	28	75.8	86.1	808	6	HS18U	X57352 Human 1-8U
	29	75.8	86.1	143779	8	AC136475	AC136475 Homo sapi
	30	75.8	86.1	156819	14	AP006283	AP006283 Homo sapi
	31	75.8	86.1	175416	14	AP006286	AP006286 Homo sapi
	32	75.8	86.1	187160	8	AC138230	AC138230 Homo sapi
	33	75.4	85.7	401	6	CQ695754	CQ695754 Sequence
	34	75.4	85.7	606	6	BD190296	BD190296 Elki phos
	35	75.4	85.7	621	6	BD190298	BD190298 Elki phos
	36	74.4	84.5	1933	6	E01979	E01979 DNA encoding
	37	73.6	83.6	422	10	AB146371	AB146371 Homo sapi
	38	73.6	83.6	84979	8	AC116049	AC116049 Homo sapi
	39	73.6	83.6	176104	14	AC016193	AC016193 Homo sapi
	40	73.6	83.6	185237	8	AC011124	AC011124 Homo sapi
	41	73.6	83.6	189572	14	AC084812	AC084812 Homo sapi
	42	72.6	82.5	47165	8	HS078113	AL121994 Human DNA
	43	72.6	82.5	204096	14	AC027704	AC027704 Homo sapi
	44	72.4	82.3	346	6	CQ661861	CQ661861 Sequence
	45	72	81.8	51649	8	BX927182	BX927182 Human DNA

#### ALIGNMENTS

BC022439 645 bp mRNA linear PRI 29-JUN-2004  
Homo sapiens interferon induced transmembrane protein 3 (1-80),  
mRNA (cDNA clone MGC:24755 IMAGE:4282809), complete cds.  
BC022439  
BC022439.1 GI:18490258  
MGC.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
1 (bases 1 to 645)

Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,  
Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Loquellano,N.A., Peters,G.J.,  
Carninci,P., Prange,C., Raha,S.S., Bosak,S.A., McGowan,P.J.,  
Abramson,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,  
McKernan,R.D., Mullahy,S.J., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,  
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,  
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,  
Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,  
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,  
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,  
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,  
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalilov,D.E.,  
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.

Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932  
2 (bases 1 to 645)

Strausberg,R.

Direct Submission

Submitted (01-FEB-2002) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 Contact: MGC help desk  
 Email: [cgapsb-remail.nih.gov](mailto:cgapsb-remail.nih.gov)  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: CLONTECH Laboratories, Inc.  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 Center, Stanford University School of Medicine, Stanford, CA 94305  
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome  
 Web site: <http://www-shgc.stanford.edu>  
 Contact: (Dickson, Mark) [mcd@paxil.stanford.edu](mailto:mcd@paxil.stanford.edu)  
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAL Plate: 35 Row: m Column: 16.

## FEATURES

source  
 1. .645  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="MGC:24755 IMAGE:4282809"  
 /tissue\_type="Brain, primitive neuroectodermal"  
 /clone\_lib="NIH\_MGC\_56"  
 /lab\_host="DH10B"  
 /note="vector: pDNR-LIB"  
 1. .645  
 /gene="IFITM3"  
 /note="synonym: 1-8U"  
 /db\_xref="GeneID:10410"  
 /db\_xref="MIM:605579"  
 51. .452  
 /gene="IFITM3"  
 /codon\_start=1  
 /product="interferon-induced transmembrane protein 3 (1-8U)"  
 /protein\_id="AAH22439.1"  
 /db\_xref="GI:18490259"  
 /db\_xref="GeneID:10410"  
 /db\_xref="MIM:605579"  
 /translation="MNQTVOTFFSPVNSGPPNVEMLKEEHEVAVLGAPHPAPPTST  
 VIHRSSTVPDHVWSLFTLWNPCLGFIAPYSKSRDRKQWGDVTGAQAYAST  
 AKCLNIWALILGILMTLLIVIPVLIFQATG"

## gene

## CDS

## ORIGIN

Query Match 92.7%; Score 81.6; DB 8; Length 645;  
 Best Local Similarity 95.5%; Pred. No. 6.6e-17;  
 Matches 84; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCCTCACAGCTTTTCTACAATGGCATTCAAT 60  
 Db 535 CGGAGCGGAGTCCTGTATCAGCCCTTTATCCTCACAGCTTTTCTACAATGGCATTCAAT 594  
 Qy 61 AAAGTGCACGTGTTCTTGGTACAAAAA 88  
 Db 595 AAAGTGCACGTGTTCTTGGTAAAAAAA 622

## RESULT 2

BV167599/c  
 LOCUS sqm5991 Human DNA (Sequenom) Homo sapiens STS 10-JUN-2004  
 DEFINITION tagged site.

ACCESSION BV167599  
 VERSION BV167599.1 GI:48001170  
 KEYWORDS STS.  
 SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Hominidae; Homo.

## REFERENCE

1 (bases 1 to 251)  
 Nelson, R.M., Marnellos, G., Kammerer, S., Hoyal, C.R., Shi, M.M.,  
 Cantor, C.R. and Braun, A.  
 Large-Scale Validation of Single Nucleotide Polymorphisms in Gene  
 Regions  
 Genome Res. (2004) In press

## JOURNAL

## COMMENT

Contact: Andreas Braun  
 Pharmaceuticals division  
 Sequenom, Inc.  
 3595 John Hopkins Court, San Diego, CA 92121, USA  
 Tel: 18582029018  
 Fax: 18582029020  
 Email: [abraham@sequenom.com](mailto:abraham@sequenom.com)  
 Primer A: No primer sequence submitted  
 Primer B: No primer sequence submitted  
 STS size: 251.

## FEATURES

source  
 1. .251  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /clone\_lib="Human DNA (Sequenom)"  
 <1. .>251

## ORIGIN

Query Match 90.9%; Score 80; DB 10; Length 251;  
 Best Local Similarity 94.3%; Pred. No. 2.5e-16;  
 Matches 83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCCTCACAGCTTTTCTACAATGGCATTCAAT 60  
 Db 107 CGGAGCGGAGTCCTGTATCAGCCCTTTATCCTCACAGCTTTTCTACAATGGCATTCAAT 48

Qy 61 AAAGTGCACGTGTTCTTGGTACAAAAA 88

Db 47 AAAGTGCACGTGTTCTTGGWAAAAAAA 20

## RESULT 3

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## Homo sapiens (human)

## Homo sapiens

## Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

## Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

## Hominidae; Homo.

## 1

## Lillie, J., Xu, Y., Wang, Y. and Steinmann, K.

## Novel genes, compositions, kits, and methods for identification,

## assessment, prevention, and therapy of breast cancer

## Patent: WO 0151628-A 346 19-JUL-2001;

## Millennium Pharmaceuticals, Inc. (US)

## Location/Qualifiers

## 1. .467

## /organism="Homo sapiens"

## /mol\_type="unassigned DNA"

## /db\_xref="taxon:9606"

## Query Match 90.9%; Score 80; DB 6; Length 467;

## Best Local Similarity 94.3%; Pred. No. 2.4e-16;

## Matches 83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

## Qy 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCCTCACAGCTTTTCTACAATGGCATTCAAT 60

## Db 67 CGGAGCGGAGTCCTGTATCAGCCCTTTATCCTCACAGCTTTTCTACAATGGCATTCAAT 126

## Qy 61 AAAGTGCACGTGTTCTTGGTACAAAAA 88



Db	127	CGAGCCGAGTCTGTATCATGCCCTTTATCTCTACACGCTTTTCTACATGGCAATTCAT	34	CGAGCCGAGTCTGTATCATGCCCTTTATCTCTACACGCTTTTCTACATGGCAATTCAT	33
Qy	61	AAAGTGCACGTGTTTCTTGTGACAAAAA	88	AAAGTGCACGTGTTTCTTGTGACAAAAA	88
Db	94	AAAGTGCACGTGTTTCTTGTGACAAAAA	121	AAAGTGCACGTGTTTCTTGTGACAAAAA	121
RESULT 6	BC070243	536 bp	DNA	630 bp	mRNA
LOCUS	Sequence 18105 from Patent WO0151628.	linear	PAT 28-JAN-2004	Homo sapiens interferon induced transmembrane protein 3 (1-8U),	linear PRI 30-JUN-2004
DEFINITION	CQ433071			mRNA (cDNA clone MGC:88228 IMAGE:30401041), complete cds.	
ACCESSION	CQ433071.1	GI:41385300		BC070243.1	GI:47692386
VERSION					
KEYWORDS					
SOURCE	Homo sapiens (human)			Homo sapiens (human)	
ORGANISM	Homo sapiens			Homo sapiens	
REFERENCE	1	Lillie, J., Xu, Y., Wang, Y. and Steinmann, K.		1	(bases 1 to 630)
AUTHORS		Novel genes, compositions, kits, and methods for identification,			
TITLE		assessment, prevention, and therapy of breast cancer			
JOURNAL		Patent: WO 0151628-A 18105 19-JUL-2001;			
FEATURES		Millennium Pharmaceuticals, Inc. (US)			
source		Location/Qualifiers			
ORIGIN		1..536			
Query Match		90.9%; Score 80; DB 6; Length 536;			
Best Local Similarity		94.3%; Pred. No. 2.3e-16;			
Matches		83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;			
Qy	1	CGAGTGCAGTCTGTATCATGCCCTTTATCTCTACACGCTTTTCTACATGGCAATTCAT	60		
Db	30	CGAGCCGAGTCTGTATCATGCCCTTTATCTCTACACGCTTTTCTACATGGCAATTCAT	89		
Qy	61	AAAGTGCACGTGTTTCTTGTGACAAAAA	88		
Db	90	AAAGTGCACGTGTTTCTTGTGACAAAAA	117		
RESULT 5	CQ425463	541 bp	DNA	630 bp	mRNA
LOCUS	Sequence 10497 from Patent WO0151628.	linear	PAT 28-JAN-2004	Homo sapiens interferon induced transmembrane protein 3 (1-8U),	linear PRI 30-JUN-2004
DEFINITION	CQ425463			mRNA (cDNA clone MGC:88228 IMAGE:30401041), complete cds.	
ACCESSION	CQ425463.1	GI:41377692		BC070243.1	GI:47692386
VERSION					
KEYWORDS					
SOURCE	Homo sapiens (human)			Homo sapiens (human)	
ORGANISM	Homo sapiens			Homo sapiens	
REFERENCE	1	Lillie, J., Xu, Y., Wang, Y. and Steinmann, K.		1	(bases 1 to 630)
AUTHORS		Novel genes, compositions, kits, and methods for identification,			
TITLE		assessment, prevention, and therapy of breast cancer			
JOURNAL		Patent: WO 0151628-A 10497 19-JUL-2001;			
FEATURES		Millennium Pharmaceuticals, Inc. (US)			
source		Location/Qualifiers			
ORIGIN		1..541			
Query Match		90.9%; Score 80; DB 6; Length 541;			
Best Local Similarity		94.3%; Pred. No. 2.3e-16;			
Matches		83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;			
Qy	1	CGAGTGCAGTCTGTATCATGCCCTTTATCTCTACACGCTTTTCTACATGGCAATTCAT	60		

/c1one="MGC:88228 IMAGE:30401041"  
/tissue\_type="Blood vessels- blood vessel, aorta, basilar  
artery."  
/clone\_lib="NIH\_MGC\_187"  
/lab\_host="DH10B"  
/note="Vector: pDNR-LIB"  
1. .630  
/gene="IFITM3"  
/note="synonym: 1-8U"  
/db\_xref="GeneID:10410"  
/db\_xref="MIM:605579"  
36. .437  
/gene="IFITM3"  
/codon\_start=1  
/product="interferon-induced transmembrane protein 3  
(1-8U)"  
/protein\_id="AAH70243.1"  
/db\_xref="GI:47682387"  
/db\_xref="GeneID:10410"  
/db\_xref="MIM:605579"  
/translation="MNTVQTFSPVNSGPPNYEMLKESHEVAVLGAHPNPPTST  
VIHIRSTVPDHVWSLENTLFNPPCLGFIAFAYSVKSRDRRMVDVGAQAYAST  
AKCLNIWALILGILMTILLIPIVLIPOAYG"

ORIGIN

Query Match 90.9%; Score 80; DB 8; Length 630;  
Best Local Similarity 94.3%; Pred. No. 2.3e-16;  
Matches 83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
  
Qy 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCCTCACAGCTTTTCTACAATGGCATTCAAT 60  
|||||  
Db 520 CGGAGCGAGTCCTGTATCAGCCCTTTATCCTCACAGCTTTTCTACAATGGCATTCAAT 579  
|||||  
  
Qy 61 AAAGTGCACGTGTTCTTGTGTACAAAAA 88  
|||||  
Db 580 AAAGTGCACGTGTTCTTGTGTACAAAAA 607  
|||||

RESULT 7

BD237072  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

BD237072 642 bp DNA linear PAT 17-JUL-2003  
Compounds for remedy and diagnosis of lung cancer and method for  
using the same.  
BD237072  
BD237072.1 GI:33046842  
JP 2002516659-A/73.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

OS Homo sapiens (human)  
PN JP 2002516659-A/73  
PD 11-JUN-2002  
PF 26-JAN-1999 JP 2000529432  
PR 28-JAN-1998 US 09/015029,28-JAN-1998 US 09/015022 PR  
18-MAR-1998 US 09/040828,18-MAR-1998 US 09/040831 PR  
23-JUL-1998 US 09/122192,23-JUL-1998 US 09/122191 PR  
22-DEC-1998 US 09/219245  
PI STEVEN G REED,MICHAEL J LODES,TONY N FRUDAKIS,RAODOH MOHAMATH  
PC C12N15/09,A61K35/14,A61K38/00,A61K39/00,A61K39/39,A61K39/395,  
PC A61K39/395,  
PC A61F11/00,A61P35/00,C07K14/47,C07K16/18,C07K19/00,C12N1/19, PC  
C12N1/21,  
PC C12N5/10,C12P21/08,C12Q1/68,G01N33/53,G01N33/574,G01N33/577//  
PC (C12N1/21,C12R1:19),C12N15/00,A61K37/02,C12N5/00 CC  
Compounds for remedy and diagnosis of lung cancer and method CC

CC same for using the  
FH key Location/Qualifiers  
FT source 1. .642  
/organism="Homo sapiens (human)".  
FEATURES  
source Location/Qualifiers  
1. .642  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

ORIGIN

Query Match 90.9%; Score 80; DB 6; Length 642;  
Best Local Similarity 94.3%; Pred. No. 2.3e-16;  
Matches 83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
  
Qy 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCCTCACAGCTTTTCTACAATGGCATTCAAT 60  
|||||  
Db 535 CGGAGCGAGTCCTGTATCAGCCCTTTATCCTCACAGCTTTTCTACAATGGCATTCAAT 594  
|||||  
  
Qy 61 AAAGTGCACGTGTTCTTGTGTACAAAAA 88  
|||||  
Db 595 AAAGTGCACGTGTTCTTGTGTACAAAAA 622  
|||||

RESULT 8

AR225472  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

AR225472 642 bp DNA linear PAT 20-DEC-2002  
Sequence 119 from patent US 6444425.  
AR225472  
AR225472.1 GI:27263418  
Unknown.  
Unclassified.  
REFERENCE  
1 (bases 1 to 642)  
Reed,S.G., Lodes,M.J., Mohamath,R. and Secrist,H.  
Compounds for therapy and diagnosis of lung cancer and methods for  
their use  
Patent: US 6444425-A 119 03-SEP-2002;  
Corixa Corporation; Seattle, WA;  
WOX;

FEATURES

source Location/Qualifiers  
1. .642  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN

Query Match 90.9%; Score 80; DB 6; Length 642;  
Best Local Similarity 94.3%; Pred. No. 2.3e-16;  
Matches 83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
  
Qy 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCCTCACAGCTTTTCTACAATGGCATTCAAT 60  
|||||  
Db 535 CGGAGCGAGTCCTGTATCAGCCCTTTATCCTCACAGCTTTTCTACAATGGCATTCAAT 594  
|||||  
  
Qy 61 AAAGTGCACGTGTTCTTGTGTACAAAAA 88  
|||||  
Db 595 AAAGTGCACGTGTTCTTGTGTACAAAAA 622  
|||||

RESULT 9

AR562895  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

AR562895 642 bp DNA linear PAT 08-OCT-2004  
Sequence 119 from patent US 6759508.  
AR562895  
AR562895.1 GI:53977724  
Unknown.  
Unclassified.  
REFERENCE  
1 (bases 1 to 642)  
Lodes,M.J., Mohamath,R., Henderson,R.A., Benson,D.R. and Secrist,H.

```

TITLE      Compositions and methods for the therapy and diagnosis of lung
JOURNAL    Cancer
FEATURES   Patent: US 679508-A 119 06-JUL-2004;
           Corixa Corporation; Seattle, WA
           Location/Qualifiers
           1..642
           /organism="unknown"
           /mol_type="genomic DNA"

ORIGIN
Query Match      90.9%; Score 80; DB 6; Length 642;
Best Local Similarity 94.3%; Pred. No. 2.3e-16;
Matches 83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy      1 CGGAGTCGAGTCCTGATCAGCCCTTTATCTCCTCACAGCCTTTCTACAAATGGCATTCAAT 60
      |||||
Db      535 CGGAGCCGAGTCCTGATCAGCCCTTTATCTCCTCACAGCCTTTCTACAAATGGCATTCAAT 594

Qy      61 AAAGTGCACGTGTTCTTCTGGTACAAAA 88
      |||||
Db      595 AAAGTGCACGTGTTCTTCTGGTAAAAAAA 622

RESULT 10
AX321588
LOCUS      AX321588          642 bp          DNA          linear          PAT 15-DEC-2001
DEFINITION Sequence 119 from Patent WO0172295.
ACCESSION AX321588
VERSION   AX321588.1 GI:17905704
KEYWORDS
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
           Hominoidea; Homo.

REFERENCE
1
AUTHORS   Reed,S.G., Lodes,M.J., Mohamath,R., Secrist,H., Benson,D.R.,
           Indrias,C.Y., Henderson,R.A., Fling,S.P., Algate,P.A., Elliot,M.,
           Mannion,J. and Kalos,M.D.
TITLE     Compositions and methods for the therapy and diagnosis of lung
JOURNAL  cancer
FEATURES  Patent: WO 0172295-A 119 04-OCT-2001;
           CORIXA CORPORATION (US)
           Location/Qualifiers
           1..642
           /organism="Homo sapiens"
           /mol_type="unassigned DNA"
           /db_xref="taxon:9606"

ORIGIN
Query Match      90.9%; Score 80; DB 6; Length 642;
Best Local Similarity 94.3%; Pred. No. 2.3e-16;
Matches 83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy      1 CGGAGTCGAGTCCTGATCAGCCCTTTATCTCCTCACAGCCTTTCTACAAATGGCATTCAAT 60
      |||||
Db      535 CGGAGCCGAGTCCTGATCAGCCCTTTATCTCCTCACAGCCTTTCTACAAATGGCATTCAAT 594

Qy      61 AAAGTGCACGTGTTCTTCTGGTACAAAA 88
      |||||
Db      595 AAAGTGCACGTGTTCTTCTGGTAAAAAAA 622

RESULT 11
AC144988/c
LOCUS      AC144988          222085 bp          DNA          linear          HTG 10-JUL-2003
DEFINITION Gorilla gorilla clone CH255-209M14, WORKING DRAFT SEQUENCE,
           8 ordered pieces.
ACCESSION AC144988
VERSION   AC144988.2 GI:32490606
KEYWORDS  HTG; HTGS, PHASE2; HTGS DRAFT.
SOURCE    Gorilla gorilla gorilla (lowland gorilla)
ORGANISM  Gorilla gorilla gorilla

```

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Gorilla.
1 (bases 1 to 222085)
Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S., Carriaga,K.,
Chu,G., Coleman,B., Coleman,H., Engle,J., Granite,S., Guan,X.,
Gupta,J., Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Hu,P.,
Hurle,B., Idol,J.R., Karlins,E., Kwong,P., Laric,P., Lee-Lin,S.-Q.,
Legaspi,R., Maduro,Q.L., Maduro,V.B., Margulies,E.H., Masiello,C.,
Maskeri,B., McDowell,J., Paguirigan,C., Pearson,R., Portnoy,M.E.,
Prasad,A., Reddix-Dugue,N., Schandler,K., Schueler,M.G., Shah,K.,
Sison,C., Stantripp,S., Thomas,J.W., Thomas,P.J., Tsipouri,V.,
Vogt,J.L., Wetherby,K.D., Wiggins,L., Young,A. and Green,E.D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 222085)
Green,E.D.
Direct Submission
Submitted (28-MAY-2003) NIH Intramural Sequencing Center, 8717
Grovemont Circle, Gaithersburg, MD 20877, USA
3 (bases 1 to 222085)
Green,E.D.
Direct Submission
Submitted (10-JUL-2003) NIH Intramural Sequencing Center, 8717
Grovemont Circle, Gaithersburg, MD 20877, USA
On Jul 10, 2003 this sequence version replaced gi:31088406.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.zoo@nhgri.nih.gov
----- Project Information
Center project name: enc
Center clone name: 209M14

```

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics

```

Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 220187 bases at least Q40
Consensus quality: 220809 bases at least Q30
Consensus quality: 221225 bases at least Q20
Insert size: 203000; agarose-fp
Insert size: 221385; sum-of-contigs
Quality coverage: 10.43x in Q20 bases; agarose-fp
Quality coverage: 9.56x in Q20 bases; sum-of-contigs
-----

```

\* NOTE: This is a 'working draft' sequence. It currently consists of 8 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

\* This sequence will be replaced

\* by the finished sequence as soon as it is available and

\* the accession number will be preserved.

\* 1 2032: contig of 2032 bp in length

\* 2033 2132: gap of unknown length

\* 2133 40439: contig of 38307 bp in length

\* 40440 40539: gap of unknown length

```

* 40540 59564: contig of 19025 bp in length
* 59565 59664: gap of unknown length
* 59666 93277: contig of 33613 bp in length
* 93278 93377: gap of unknown length
* 93378 154966: contig of 61589 bp in length
* 154967 155066: gap of unknown length
* 155067 168650: contig of 13584 bp in length
* 168651 168750: gap of unknown length
* 168751 185207: contig of 16457 bp in length
* 185208 185307: gap of unknown length
* 185308 222085: contig of 36778 bp in length.
FEATURES
    source
        1..222085
            /organism="Gorilla gorilla gorilla"
            /mol_type="genomic DNA"
            /sub_species="gorilla"
            /db_xref="taxon:9595"
            /clone="CH255-209M14"
            /clone_lib="CH255"
        1..2032
            /note="assembly_fragment
            missing 17 clone end on 5' end of insert"
            2033..2132
            /estimated_length=unknown
            /note="assembly_fragment"
            2133..40439
            /note="assembly_fragment"
            40440..40539
            /estimated_length=unknown
            40540..59564
            /note="assembly_fragment"
            59565..59664
            /estimated_length=unknown
            59665..93277
            /note="assembly_fragment"
            93278..93377
            /estimated_length=unknown
            93378..154966
            /note="assembly_fragment"
            154967..155066
            /estimated_length=unknown
            155067..168650
            /note="assembly_fragment"
            168651..168750
            /estimated_length=unknown
            168751..185207
            /note="assembly_fragment"
            185208..185307
            /estimated_length=unknown
            185308..222085
            /note="assembly_fragment
            clone_end:SP6
            vector_side:right"
ORIGIN
Query Match          90.9%; Score 80; DB 14; Length 222085;
Best Local Similarity 94.3%; Pred. No. 1.5e-16;
Matches 83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 1 CGGAGTCGAGTCCTGATCAGCCCTTTATCCTCACAGCGTTTTCACAATGGCATTCAAT 60
    |||||
Db 116410 CGGAGCCGAGTCCTGATCAGCCCTTTATCCTCACAGCGTTTTCACAATGGCATTCAAT 116351
    |||||
Qy 61 AAATGACGCTGTTCTTCTGGTACAAAAA 88
    |||||
Db 116350 AAATGACGCTGTTCTTCTGGTACAAAAA 116323
    |||||
RESULT 12
BC006794
LOCUS
DEFINITION Homo sapiens interferon induced transmembrane protein 3 (1-8U),
            mRNA (cdna clone MGC:5225 IMAGE:2986145), complete cds.
ACCESSION BC006794

```

```

VERSION
KEYWORDS
SOURCE
    ORGANISM
        Homo sapiens (human)
        Homo sapiens
        Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
        Homnidae; Homo.
REFERENCE
    1 (bases 1 to 637)
        Strausberg RL, Feingold EA, Grouse LH, Derge JG, Klausner RD,
        Collins FS, Wagner L, Shenmen CW, Schuler GD, Altschul SF, Zeeberg
        B, Buetow KH, Schaefer CF, Bhat NK, Hopkins RF, Jordan H, Moore T,
        Max SI, Wang J, Haieh F, Diatchenko L, Marusina K, Farmer AA, Rubin
        GM, Hong L, Stapleton M, Soares MB, Bonaldo MF, Casavant TL,
        Schetz TE, Brownstein MJ, Uedin TB, Toshiyuki S, Carninci P,
        Prange C, Rana SS, Loquellano NA, Peters GJ, Abramson RD, Mullahy
        SJ, Bosak SA, McEwan PJ, McKernan KJ, Malek JA, Gunaratne PH,
        Richards S, Worley KC, Hale S, Garcia AM, Gay LJ, Hulyk SW,
        Villalon DK, Muzny DM, Sodergren EJ, Lu X, Gibbs RA, Fahey J,
        Helton E, Kettaman M, Madan A, Rodriguez S, Sanchez A, Whiting M,
        Madan A, Young AC, Shevchenko Y, Bouffard GG, Blakealey RW,
        Touchman JW, Green ED, Dickson MC, Rodriguez AC, Grimwood J, U,
        Smalutz J, Myers RM, Butterfield YS, Krzywinski MI, Skalska U,
        Smailus DE, Schnerch A, Schein JE, Jones SJ and Marra MA.
        Mammalian Gene Collection Program Team
        Generation and initial analysis of more than 15,000 full-length
        human and mouse cDNA sequences
        Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
        12477932
    2 (bases 1 to 637)
        Director MGC Project.
        Direct Submission
        Submitted (27-APR-2001) National Institutes of Health, Mammalian
        Gene Collection (MGC), Cancer Genomics Office, National Cancer
        Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
        USA
        NIH-MGC Project URL: http://mgc.nci.nih.gov
        Contact: MGC help desk
        Email: cgapbs-x@mail.nih.gov
        Tissue Procurement: ATCC
        CDNA Library Preparation: Life Technologies, Inc.
        CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
        DNA Sequencing by: Baylor College of Medicine Human Genome
        Sequencing Center
        Center code: BCM-HGSC
        Web site: http://www.hgsc.bcm.tmc.edu/cdna/
        Contact: amg@bcm.tmc.edu
        Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,
        Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
        A.N., Gibbs, R.A.
        Clone distribution: MGC clone distribution information can be found
        through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
        Series: IRAX Plate: 3 Row: j Column: 11
        This clone was selected for full length sequencing because it
        passed the following selection criteria: matched mRNA gi: 11995467.
FEATURES
    source
        1..637
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="MGC:5225 IMAGE:2986145"
            /tissue_type="Cervix, carcinoma"
            /clone_lib="NIH_MGC_12"
            /lab_host="DH10B"
            /note="Vector: pCMV-SPORT6"
            1..637
            /gene="IFITM3"
            /note="synonym: 1-8U"
            /db_xref="GeneID:10410"
            /db_xref="MIM:605579"
            60..461
            /gene="IFITM3"
            /codon_start=1
gene
CDS

```

```
/product="interferon-induced transmembrane protein 3
(1-8U)"
/protein_id="AAH06794.1"
/db_xref="GI:13905022"
/db_xref="GeneID:10410"
/db_xref="MIM:605579"
/translation="MNHTVQTFFSPVNSGQPPNVEMLKEHEVAVLGAPHNPAPPTST
VHIRSETSPDHVWVSLFNTLFWNPCCLGFIAPFAYSVKSRDRKMGVDVTGAQAYAST
AKCLNIWALILGILMTILLIIVPLIFQYG"

ORIGIN
Query Match      89.1%; Score 78.4; DB 8; Length 637;
Best Local Similarity 93.2%; Pred. No. 8.1e-16;
Matches 82; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CGGAGTCGAGTCTGTATCAGCGCTTTATCCTCACACGCTTTTCTACAAATGCGATTCAAT 60
|||||
Db 544 CGGAGCCGAGTCTGTATCAGCGCTTTATCCTCACACGCTTTTCTACAAATGCGATTCAAT 603
|||||

QY 61 AAGTGCACGTGTTCTTGGTACAAAA 88
|||||
Db 604 AAGTGCACGTGTTCTGAAAAAATAAA 631
|||||

RESULT 13
AX281855 695 bp DNA linear PAT 02-NOV-2001
LOCUS
DEFINITION Sequence 264 from Patent WO0177389.
ACCESSION AX281855
VERSION AX281855.1 GI:16609106
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE
1 Shiffman,D., Somogyi,R., Lawn,R., Seilhamer,J.J., Porter,G.J.,
Mikita,T. and Tai,J
Genes expressed in foam cell differentiation
Patent: WO 0177389-A 264 18-OCT-2001;
Incyte Genomics, Inc. (US)
Location/Qualifiers
1..695
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 444902.6"

ORIGIN
Query Match      89.1%; Score 78.4; DB 6; Length 695;
Best Local Similarity 93.2%; Pred. No. 8.1e-16;
Matches 82; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CGGAGTCGAGTCTGTATCAGCGCTTTATCCTCACACGCTTTTCTACAAATGCGATTCAAT 60
|||||
Db 585 CGGAGCCGAGTCTGTATCAGCGCTTTATCCTCACACGCTTTTCTACAAATGCGATTCAAT 644
|||||

QY 61 AAGTGCACGTGTTCTTGGTACAAAA 88
|||||
Db 645 AAGTGCACGTGTTCTGAAAAAATAAA 672
|||||

RESULT 14
BC008417 749 bp mRNA linear PRI 29-JUN-2004
LOCUS
DEFINITION Homo sapiens interferon induced transmembrane protein 3 (1-8U),
mRNA (cDNA clone MGC:14565 IMAGE:4075453), complete cds.
ACCESSION BC008417
VERSION BC008417.1 GI:14250037
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
```

```
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 749)
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Haieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.P., Casavant,T.L.,
Schetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Pahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Green,E.D.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Schmutz,J.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
Scherch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 749)
Strausberg,R.
Direct Submission
Submitted (25-MAY-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www-shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 21 Row: 0 Column: 15
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, Similarity but not identity to protein.
Location/Qualifiers
1..749
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:14565 IMAGE:4075453"
/tissue_type="Brain, glioblastoma"
/clone_lib="NIH MGC_57"
/lab_host="DH10B"
/note="Vector: pDNR-LIB"
1..749
/gene="IFITM3"
/note="synonym: 1-8U"
/db_xref="GeneID:10410"
/db_xref="MIM:605579"
45..446
/gene="IFITM3"
/codon_start=1
/product="interferon-induced transmembrane protein 3
(1-8U)"
/protein_id="AAH08417.1"

FEATURES
source
```

```
/db_xref="GI:14250038"
/db_xref="GeneID:10410"
/db_xref="WIM:605579"
/translation="MQGTVTFFSPVNSGQPPNYEMLKEHEVAVILGAPHPAPPTST
VIHRSSTVDFDHWVSLNLFMNPCCLGFIAPAYSVKSRDRRMVGDVTGAQAYAST
AKCLNIWALILGILMTILLIVIPVIFQAYG"

ORIGIN
Query Match      89.1%; Score 78.4; DB 8; Length 749;
Best Local Similarity 93.2%; Pred. No. 8e-16;
Matches 82; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCTCACACGCTTTTCTACAATGGCATTCAAT 60
    |||||
Db 529 CGGAGCCGAGTCCTGTATCAGCCCTTTATCTCACACGCTTTTCTACAATGGCATTCAAT 588
    |||||

Qy 61 AAAGTGACGCTGTTTCTTGGTACAAAA 98
    |||||
Db 589 AAAGTGACGCTGTTTCTGAAAAA 616

RESULT 15
BD222194
LOCUS Human nucleic acid sequence originating in normal uterine tissue.
DEFINITION
ACCESSION BD222194
VERSION BD222194.1 GI:33031964
KEYWORDS JP 2002512017-A/23.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 451)
AUTHORS Specht,T., Hinzmann,B., Schmitt,A., Pilarsky,C., Dahl,E. and
Rosenthal,A.
TITLE Human nucleic acid sequence originating in normal uterine tissue
JOURNAL Patent: JP 2002512017-A 23 23-APR-2002,
METAGEN GESELLSCHAFT FUER GENOM FORSCHUNG MBH
COMMENT OS Homo sapiens (human)
PN JP 2002512017-A/23
PD 23-APR-2002
PF 15-APR-1999 JP 2000544691
PR 17-APR-1998 DE 198 17 946.4
PI THOMAS SPECHT,BERND HINZMANN,ARMIN SCHMITT,CHRISTIAN PILARSKY,
PI EDGAR DAHL,
PI ANDRE ROSENTHAL
PC C12N15/09,A61K38/00,A61K48/00,A61P35/00,C07K14/47,C07K16/18,
PC C12N1/19,
PC C12N1/21,C12N5/00,C12P21/08,C12Q1/68,G01N33/68,C12N15/00, PC
A61K37/02,
PC C12N5/00
CC Human nucleic acid sequence originating in normal uterine CC
tissue
FH Key Location/Qualifiers
FT source 1..451
FT /organism="Homo sapiens (human)".

FEATURES
source Location/Qualifiers
1..451
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match      87.3%; Score 76.8; DB 6; Length 451;
Best Local Similarity 92.0%; Pred. No. 2.9e-15;
Matches 81; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCTCACACGCTTTTCTACAATGGCATTCAAT 60
    |||||
Db 356 CGGAGCCGAGTCCTGTATCAGCCCTTTATCTCACACGCTTTTCTACAATGGCATTCAAT 415
    |||||

Qy 61 AAAGTGACGCTGTTTCTTGGTACAAAA 88
```

Db 416 AAAGTGCAAGTGTTTCTGGTGA 443

Search completed: December 7, 2005, 04:22:01  
Job time : 781.261 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 11:47:47 ; Search time 138.343 Seconds  
(without alignments)  
4239.411 Million cell updates/sec

Title: US-09-980-046B-5  
Perfect score: 88  
Sequence: 1 cggagtcgagtcctgtatca.....cgtgttcttgggtacaaaaa 88

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_21.\*  
1: geneseqn1980s.\*  
2: geneseqn1990s.\*  
3: geneseqn2000s.\*  
4: geneseqn2001as.\*  
5: geneseqn2001bs.\*  
6: geneseqn2002as.\*  
7: geneseqn2002bs.\*  
8: geneseqn2003as.\*  
9: geneseqn2003bs.\*  
10: geneseqn2003cs.\*  
11: geneseqn2003ds.\*  
12: geneseqn2004as.\*  
13: geneseqn2004bs.\*  
14: geneseqn2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	88	100.0	88	4 AAC89736	Aac89736 Human gas
2	80	90.9	467	4 AAL07889	Aal07889 Human bre
3	80	90.9	483	9 ACH33912	Ach33912 Human end
4	80	90.9	536	4 AAL25635	Aal25635 Human bre
5	80	90.9	541	4 AAL18030	Aal18030 Human bre
6	80	90.9	561	10 ADK61256	Adk61256 Ovarian c
7	80	90.9	623	5 AAS75777	Aas75777 DNA encod
8	80	90.9	642	2 AAZ07228	Aaz07228 Human lun
9	80	90.9	642	3 AAC79118	Aac79118 Human lun
10	80	90.9	642	4 AAD23193	Aad23193 Human lun
11	80	90.9	642	10 ADD66427	Add66427 Human lun
12	80	90.9	642	10 ADE87681	Ade87681 Human lun
13	80	90.9	993	11 ACN88677	Acn88677 Breast ca
14	78.4	89.1	682	10 ADJ56210	Adj56210 Human cdn
15	78.4	89.1	695	6 AAS95009	Aas95009 Human DNA
16	76.8	87.3	451	2 AAZ41347	Aaz41347 Human nor
17	76	86.4	1408	12 ADP83378	Adp83378 Breast sp
18	76	86.4	1889	12 ADP83377	Adp83377 Breast sp
19	75.8	86.1	118	12 ADL83354	Adl83354 Human ost

20	75.8	86.1	167	2 AAT21450	Aat21450 Human gen
21	75.8	86.1	594	13 ADU14728	Adu14728 Solid tum
c 22	75.8	86.1	668	4 AAL26695	Aal26695 Human bre
23	75.8	86.1	808	3 AAA46668	Aaa46668 cDNA of I
24	75.8	86.1	808	8 ACC42318	Acc42318 Human MAP
25	75.8	86.1	808	10 ADK61255	Adk61255 Ovarian c
26	75.8	86.1	808	11 ADI31953	Adi31953 Human CDN
27	75.8	86.1	808	13 ADS84020	Ads84020 Human lym
28	75.8	86.1	808	14 ADX08047	Adx08047 Cyclin-de
29	75.8	86.1	808	14 ADY20076	Ady20076 DNA encod
30	75.8	86.1	808	14 ADY86641	Ady86641 Human int
31	75.8	86.1	808	14 AEA23687	Aea23687 Human PRO
32	75.4	85.7	606	8 ACC42317	Acc42317 Human MAP
33	75.4	85.7	621	8 ACC42319	Acc42319 Human MAP
34	74.8	85.0	545	10 ADK61257	Adk61257 Ovarian c
c 35	74.4	84.5	1933	1 AAN90120	Aan90120 Human lip
36	71	80.7	600	13 ADT50791	Adt50791 Cancer re
37	70.6	80.2	627	5 AAS75776	Aas75776 DNA encod
38	70.6	80.2	678	12 ADL07643	Adl07643 Human 1-8
39	70.6	80.2	1016	5 AAS83009	Aas83009 DNA encod
c 40	70.4	80.0	283	4 AAL20125	Aal20125 Human bre
41	70.2	79.8	761	6 ABQ56182	Abq56182 Human ova
42	69.6	79.1	407	3 AAF18371	Aaf18371 Lung canc
c 43	69.4	78.9	245	11 ACN90202	Acn90202 Breast ca
44	69	78.4	762	5 AAS87704	Aas87704 DNA encod
45	68.8	78.2	148834	6 ABK83570	Abk83570 Human CDN

ALIGNMENTS

RESULT 1  
AAC89736  
ID AAC89736 standard; cDNA; 88 BP.  
XX  
AC AAC89736;  
XX  
12-MAR-2001 (first entry)  
XX Human gastrointestinal inflammation-related cDNA, SEQ ID NO: 5.  
XX Human; cytostatic; immunomodulator; immunostimulant; vulnery;  
XX anti-inflammatory; neuroprotective; antibacterial; gene therapy;  
XX gastrointestinal inflammation; immune system disorder; genetic disorder;  
XX cancer; autoimmune disorder; infection; wound healing; ss.  
XX Homo sapiens.  
OS  
FN WO200073324-A2.  
XX  
PD 07-DEC-2000.  
XX  
PF 01-JUN-2000; 2000WO-US015191.  
XX  
PR 01-JUN-1999; 99US-0137058P.  
XX (DIGI-) DIGITAL GENE TECHNOLOGIES INC.  
XX Youakim A, Dubose RF, Sims JE, Pribyl TM, Hillbush BS, Hasel KW;  
XX WPI; 2001-061508/07.  
XX  
XX New polynucleotides and polypeptides, useful in gene therapy and in  
PT diagnosing a pathological condition, e.g. for modulating gene expression  
PT in gastrointestinal inflammation, or for treating cancers or genetic  
XX disorders.  
PS Claim 1; Page 85; 108pp; English.  
XX  
XX The present sequence is one of a number of isolated human polynucleotides  
CC which are useful in gene therapy, and for diagnosing a pathological  
CC condition or a susceptibility to it. In particular, the polynucleotides  
CC are useful for modulating gene expression in gastrointestinal





CC responsible for genetic disorders and other traits. The nucleotide  
CC sequences are also useful as hybridisation probes, as oligomers for PCR,  
CC for chromosome and gene mapping, in the recombinant production of  
CC protein, or in generating antisense DNA or RNA. The purified polypeptide  
CC is useful for generating antibodies specific for it. The present sequence  
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data  
CC for this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from USPTO at  
CC seqdata.uspto.gov/sequence.html?DocID=20030073623  
XX

SQ Sequence 483 BP; 121 A; 138 C; 112 G; 110 T; 0 U; 2 Other;

Query Match 90.9%; Score 80; DB 9; Length 483;  
Best Local Similarity 94.3%; Pred. No. 7.3e-19;  
Matches 83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCTCTCACAGCTTTTCTACAATGGCATTCAAT 60  
Db 217 CGGAGCCGAGTCCTGTATCAGCCCTTTATCTCTCACAGCTTTTCTACAATGGCATTCAAT 276

Qy 61 AAAGTGACGCTGTTCTTGGTACAAAA 88  
Db 277 AAAGTGACGCTGTTCTTGGTACAAAA 304

## RESULT 4

AAL25635  
ID AAL25635 standard; cDNA; 536 BP.

XX  
AC AAL25635;

XX  
DT 07-DEC-2001 (first entry)

XX Human breast cancer expressed polynucleotide 18092.

XX Human; breast cancer; cell marker; cytostatic; ss.

XX Homo sapiens.

XX WO200151628-A2.

XX 19-JUL-2001.

XX 10-JAN-2001; 2001WO-US000798.

XX 14-JAN-2000; 2000US-0176077P.

XX 14-MAR-2000; 2000US-0189167P.

XX 24-MAR-2000; 2000US-0192099P.

XX 29-MAR-2000; 2000US-0193480P.

XX 15-MAY-2000; 2000US-0205230P.

XX 09-JUN-2000; 2000US-0211315P.

XX 25-JUL-2000; 2000US-0220534P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Lillie J, Xu Y, Wang Y, Steinmann K;

XX WPI; 2001-451856/48.

XX New peptide useful as a marker for the diagnosis of breast cancer.

XX Claim 1; Page 3344; 3695pp; English.

XX The invention relates to human breast cancer expressed polynucleotides  
(AAL07544-AAL26789) and methods of assessing whether a patient is  
CC afflicted with breast cancer by examining the correlation between the  
CC expression of certain markers and the cancerous state of breast cells.  
CC The polynucleotides and encoded polypeptides are potential markers for  
CC detecting, diagnosing, monitoring, characterising treating and  
CC potentially preventing breast cancer. The polynucleotides and encoded  
CC polypeptides are also useful for isolating compounds with cytostatic  
XX activity

SQ Sequence 536 BP; 193 A; 110 C; 65 G; 168 T; 0 U; 0 Other;

Query Match 90.9%; Score 80; DB 4; Length 536;  
Best Local Similarity 94.3%; Pred. No. 7.5e-19;  
Matches 83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCTCTCACAGCTTTTCTACAATGGCATTCAAT 60  
Db 30 CGGAGCCGAGTCCTGTATCAGCCCTTTATCTCTCACAGCTTTTCTACAATGGCATTCAAT 89

Qy 61 AAAGTGACGCTGTTCTTGGTACAAAA 88

Db 90 AAAGTGACGCTGTTCTTGGGAGAAAAA 117

## RESULT 5

AAL18030  
ID AAL18030 standard; cDNA; 541 BP.

XX  
AC AAL18030;

XX 07-DEC-2001 (first entry)

XX Human breast cancer expressed polynucleotide 10487.

XX Human; breast cancer; cell marker; cytostatic; ss.

XX Homo sapiens.

XX WO200151628-A2.

XX 19-JUL-2001.

XX 10-JAN-2001; 2001WO-US000798.

XX 14-JAN-2000; 2000US-0176077P.

XX 14-MAR-2000; 2000US-0189167P.

XX 24-MAR-2000; 2000US-0192099P.

XX 29-MAR-2000; 2000US-0193480P.

XX 15-MAY-2000; 2000US-0205230P.

XX 09-JUN-2000; 2000US-0211315P.

XX 25-JUL-2000; 2000US-0220534P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Lillie J, Xu Y, Wang Y, Steinmann K;

XX WPI; 2001-451856/48.

XX New peptide useful as a marker for the diagnosis of breast cancer.

XX Claim 1; Page 1874; 3695pp; English.

XX The invention relates to human breast cancer expressed polynucleotides  
(AAL07544-AAL26789) and methods of assessing whether a patient is  
CC afflicted with breast cancer by examining the correlation between the  
CC expression of certain markers and the cancerous state of breast cells.  
CC The polynucleotides and encoded polypeptides are potential markers for  
CC detecting, diagnosing, monitoring, characterising treating and  
CC potentially preventing breast cancer. The polynucleotides and encoded  
CC polypeptides are also useful for isolating compounds with cytostatic  
XX activity

SQ Sequence 541 BP; 194 A; 110 C; 68 G; 169 T; 0 U; 0 Other;

Query Match 90.9%; Score 80; DB 4; Length 541;  
Best Local Similarity 94.3%; Pred. No. 7.5e-19;  
Matches 83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCTCTCACAGCTTTTCTACAATGGCATTCAAT 60  
Db 34 CGGAGCCGAGTCCTGTATCAGCCCTTTATCTCTCACAGCTTTTCTACAATGGCATTCAAT 93



AAZ07228  
ID AAZ07228 standard; cDNA; 642 BP.  
AC AAZ07228;  
XX  
DT 13-OCT-1999 (first entry)  
XX  
DE Human lung tumour protein SALT-T8 5' cDNA sequence.  
XX  
KW Human; lung tumour protein; therapy; diagnosis; lung cancer; vaccine;  
KW immunotherapy; detection; inhibition; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO9938973-A2.  
XX  
XX 05-AUG-1999.  
PD  
XX 26-JAN-1999; 99WO-US001642.  
XX  
XX 28-JAN-1998; 98US-00015022.  
PR  
XX 28-JAN-1998; 98US-00015029.  
PR  
XX 18-MAR-1998; 98US-00040828.  
PR  
XX 18-MAR-1998; 98US-00040831.  
PR  
XX 23-JUL-1998; 98US-00122191.  
PR  
XX 23-JUL-1998; 98US-00122192.  
PR  
XX 22-DEC-1998; 98US-00219245.  
XX  
XX (CORI-) CORIXA CORP.  
FA  
XX Reed SG, Lodes MJ, Frudakis TN, Mohamath R;  
XX  
XX WPI; 1999-479187/40.  
DR  
XX P-PSDB; AAY29544.  
DR  
XX  
XX Lung tumor specific polynucleotides for inhibiting the development of  
PT lung cancer.  
XX  
XX Claim 13; Page 122; 171pp; English.  
XX  
XX The present invention describes lung tumour specific polynucleotides and  
CC tumour antigens. AAZ07144 to AAZ07246 and AAZ08301 to AAZ08325 represent  
CC specifically claimed polynucleotides, and AAY29486 to AAY29571 represent  
CC amino acid sequences from the present invention. The lung tumour specific  
CC polynucleotides and polypeptides can be used in pharmaceutical  
CC compositions and vaccines to inhibit the development of lung cancer. They  
CC can also be used to detect lung cancer in a patient. Probes and  
CC antibodies derived from the lung tumour sequences are useful in detection  
CC of lung cancer  
XX  
XX Sequence 642 BP; 142 A; 215 C; 147 G; 138 T; 0 U; 0 Other;  
SQ  
Query Match 90.9%; Score 80; DB 2; Length 642;  
Best Local Similarity 94.3%; Pred. No. 7.9e-19;  
Matches 83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
Qy 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCTCTCACACGCTTTCTTACAATGGCATTCAAT 60  
Db 535 CGGAGCCGAGTCCTGTATCAGCCCTTTATCTCTCACACGCTTTCTTACAATGGCATTCAAT 594  
Qy 61 AAAGTGCACGTGTTTCTTGGTACAAAAA 88  
Db 595 AAAGTGCACGTGTTTCTTGGTACAAAAA 622  
RESULT 9  
AAC79118  
ID AAC79118 standard; cDNA; 642 BP.  
XX  
AC AAC79118;  
XX  
XX 05-FEB-2001 (first entry)  
DT  
XX

DE Human lung tumour-specific cDNA #71.  
XX  
KW Lung tumour protein; lung cancer; cytostatic; vaccine; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200060077-A2.  
XX  
XX 12-OCT-2000.  
PD  
XX 30-MAR-2000; 2000WO-US008560.  
PF  
XX 02-APR-1999; 99US-00285323.  
PR  
XX 09-AUG-1999; 99US-00370838.  
PR  
XX 30-DEC-1999; 99US-00476235.  
PR  
XX 03-MAR-2000; 2000US-00518809.  
XX  
XX (CORI-) CORIXA CORP.  
FA  
XX Reed SG, Lodes MJ, Mohamath R, Secrist H;  
XX  
XX WPI; 2000-638466/61.  
DR  
XX P-PSDB; AAB44456.  
DR  
XX Novel lung tumor polypeptides and polynucleotides, useful for detecting,  
XX monitoring or treating cancer, especially lung cancer.  
XX  
XX Claim 1; Page 139; 243pp; English.  
XX  
XX The present sequence is given in a specification relating to compounds  
CC for therapy and diagnosis of lung cancer. Polypeptides comprising at  
CC least an immunogenic part of a lung tumour protein are disclosed. The  
CC polypeptides are useful for inhibiting the development of cancer,  
CC especially lung cancer. Samples of T cells expressing the polypeptides  
CC may be used to inhibit the development of cancer. The polypeptides are  
CC also useful for detecting and monitoring the progression of cancer,  
CC especially lung cancer  
XX  
XX Sequence 642 BP; 142 A; 215 C; 147 G; 138 T; 0 U; 0 Other;  
SQ  
Query Match 90.9%; Score 80; DB 3; Length 642;  
Best Local Similarity 94.3%; Pred. No. 7.9e-19;  
Matches 83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
Qy 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCTCTCACACGCTTTCTTACAATGGCATTCAAT 60  
Db 535 CGGAGCCGAGTCCTGTATCAGCCCTTTATCTCTCACACGCTTTCTTACAATGGCATTCAAT 594  
Qy 61 AAAGTGCACGTGTTTCTTGGTACAAAAA 88  
Db 595 AAAGTGCACGTGTTTCTTGGTACAAAAA 622  
RESULT 10  
AAD23193  
ID AAD23193 standard; cDNA; 642 BP.  
XX  
AC AAD23193;  
XX  
XX 26-FEB-2002 (first entry)  
DT  
XX Human lung tumour-specific protein SALT-T8 cDNA.  
XX  
XX Human; lung tumour protein; immunostimulant; cytostatic; gene therapy;  
KW antisense-therapy; vaccine; immune response; lung cancer; SALT-T8; ss.  
XX  
XX Homo sapiens.  
OS  
XX Key Location/Qualifiers  
XX CDS 51..452  
XX /\*tag= a  
XX /product= "Human SALT-T8 protein"  
XX

```

PN WO200172295-A2.
XX
PD
XX
XX 04-OCT-2001.
XX
XX 28-MAR-2001; 2001WO-US009991.
XX
XX 29-MAR-2000; 2000US-00538037.
XX
XX 05-JUN-2000; 2000US-00588937.
XX
XX 18-AUG-2000; 2000US-00640878.
XX
XX 22-SEP-2000; 2000US-0234517P.
XX
XX 01-NOV-2000; 2000US-00704512.
XX
XX 14-DEC-2000; 2000US-00738973.
XX
XX (CORI-) CORIXA CORP.
XX
XX Reed SG, Lodes MJ, Mohamath R, Secrist H, Benson DR, Indirias CY;
PI Henderson RA, Fling SP, Algate PA, Elliot M, Mannion J, Kalos MD;
XX
XX WPI; 2001-639201/73.
XX
XX P-PSDB; AAE13797.
XX
XX New human lung-specific polynucleotides and polypeptides for the
PT diagnosis and treatment of disease e.g. lung cancer.
XX
XX Example 6; Page 193; 378pp; English.
XX
XX The invention relates to isolated lung tumour-specific proteins and their
CC corresponding cDNA molecules. Lung tumour-specific proteins and their
CC antigen-presenting cells are useful for stimulating and/or expanding T
CC cells specific for a tumour protein, and for inhibiting the development
CC of cancer. The invention also relates to a composition useful for
CC stimulating an immune response, and for treating cancer. The lung tumour
CC specific oligonucleotide is useful in gene therapy and for diagnosis,
CC detection and treatment of lung cancer. The present sequence is a cDNA
CC encoding human lung tumour-specific protein
XX
XX Sequence 642 BP; 142 A; 215 C; 147 G; 138 T; 0 U; 0 Other;
SQ
Query Match 90.9%; Score 80; DB 4; Length 642;
Best Local Similarity 94.3%; Pred. No. 7.9e-19;
Matches 83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCTCCTCAGCGCTTTCTACAATGGCATTCAAT 60
DB 535 CGGAGCGAGTCCTGTATCAGCCCTTTATCTCCTCAGCGCTTTCTACAATGGCATTCAAT 594
QY 61 AAAGTGCACGTGTTCTTGGTACAAAAA 88
DB 595 AAAGTGCACGTGTTCTTGGTACAAAAA 622
RESULT 11
ADD66427
ID ADD66427 standard; cDNA; 642 BP.
XX
XX ADD66427;
XX
XX 15-JAN-2004 (first entry)
XX
XX Human lung tumour-specific related cDNA, SEQ ID No 119.
XX
XX expression control; cancer; T cell; tumour; immune; cytostatic; vaccine;
XX human; lung tumour-specific; ss.
XX
XX Homo sapiens.
XX
XX WO200292001-A2.
XX
XX 21-NOV-2002.
XX
XX 10-MAY-2002; 2002WO-US014975.
XX
XX 11-MAY-2001; 2001US-00854133.
XX
XX
XX (CORI-) CORIXA CORP.
XX
XX Lodes MJ, Wang T, Fan L, Algate PA, Mcneill PD;
XX
XX WPI; 2003-120592/11.
XX
XX New polynucleotide and polypeptide, useful for preparing a composition
PT for diagnosing, treating or preventing cancer.
XX
XX Disclosure; SEQ ID NO 119; 494pp; English.
XX
XX The invention relates to a novel isolated polynucleotide comprising one
CC of 32 47-6080 base pair sequences, given in the specification, or their
CC complements or degenerate variants, at least 20 contiguous residues of a
CC sequence in, or having at least 75 or 90 % identity with the isolated
CC polynucleotide, or that hybridise with the polynucleotide. The invention
CC further comprises: an isolated polypeptide; an expression vector
CC comprising the polynucleotide operably linked to an expression control
CC sequence; a host cell transformed or transfected with the expression
CC vector; an isolated antibody or its antigen-binding fragment that
CC specifically binds to the polypeptide; a method for detecting the
CC presence of a cancer in a patient; a fusion protein comprising the
CC polypeptide; an oligonucleotide that hybridises to the isolated
CC polynucleotide under moderately stringent conditions; a method for
CC stimulating and/or expanding T cells specific for a tumour protein; an
CC isolated T cell population; a composition comprising a first component
CC consisting of carriers and immunostimulants and a second component; a
CC method for stimulating an immune response in a patient; a method for
CC treating cancer in a patient; a method for determining cancer in a
CC patient; a diagnostic kit comprising at least one oligonucleotide or
CC antibody and a detection reagent comprising a reporter group; and a
CC method for inhibiting the development of cancer in a patient. The
CC compositions of the invention have cytostatic activity and can be used to
CC create a vaccine. The isolated polynucleotide is useful for preparing a
CC composition for diagnosing, treating or preventing cancer. This
CC polynucleotide sequence represents a human lung tumour-specific cDNA
CC sequence relating to the invention.
XX
XX Sequence 642 BP; 142 A; 215 C; 147 G; 138 T; 0 U; 0 Other;
SQ
Query Match 90.9%; Score 80; DB 10; Length 642;
Best Local Similarity 94.3%; Pred. No. 7.9e-19;
Matches 83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCTCCTCAGCGCTTTCTACAATGGCATTCAAT 60
DB 535 CGGAGCGAGTCCTGTATCAGCCCTTTATCTCCTCAGCGCTTTCTACAATGGCATTCAAT 594
QY 61 AAAGTGCACGTGTTCTTGGTACAAAAA 88
DB 595 AAAGTGCACGTGTTCTTGGTACAAAAA 622
RESULT 12
ADE87681
ID ADE87681 standard; cDNA; 642 BP.
XX
XX ADE87681;
XX
XX 29-JAN-2004 (first entry)
XX
XX Human lung tumour antigen cDNA #71.
XX
XX Human; lung tumour antigen; gene; ss; cancer; lung cancer; CD4+; CD8+;
XX T cell; immune response; immunostimulant; cytostatic.
XX
XX Homo sapiens.
XX
XX US2003118599-A1.
XX
XX 26-JUN-2003.
XX

```



```
XX 23-FEB-2001; 2001US-0270784P.
PR (STUA/) STUART S G.
XX (NUCH/) NUCHTERN J G.
PA (PLON/) PLON S E.
PA (SHOH/) SHOHET J M.
XX
PI Stuart SG, Nuchtern JG, Plon SE, Shohet JM;
XX
XX WPI; 2003-635698/60.
DR
XX
XX New genes regulated by MYCN activation, useful in gene therapy, other
PT particularly for treating a subject with e.g. neuroblastoma or other
PT cancers, or for diagnosing, staging or monitoring the treatment of the
PT cancer.
XX
XX Claim 1; SEQ ID NO 16; 27pp; English.
PS
XX
CC This invention relates to novel isolated cDNAs that are differentially
CC expressed in MYCN activated cells. Specifically, it refers to
CC polynucleotide sequences that exhibit differential expression patterns in
CC cells activated by the transactivator MYCN, where MYCN is a proto-
CC oncogene that is amplified in neuroblastoma cells and is common in small
CC cell lung cancers. The present invention describes these cDNA molecules
CC as useful for in hybridisation assays to detect expression of nucleic
CC acids (or complementary nucleic acids) in a present in a given sample, as
CC well as for screening assays by identifying molecules or compounds that
CC specifically bind the cDNA as a ligand and modulate function or activity.
CC Accordingly, these compositions exhibit cytostatic activity and can also
CC be used for gene therapy purposes. This polynucleotide sequence is a cDNA
CC that is differentially expressed in MYCN activated cells, given in an
CC exemplification of the invention. NOTE: This sequence does not appear in
CC the printed specification but has been obtained in electronic format from
CC the US Patent Office at
CC ftp.seqdata.uspto.gov/sequence.html?DocID=20030119009.
XX
XX Sequence 682 BP; 152 A; 225 C; 161 G; 144 T; 0 U; 0 Other;
SQ
Query Match 89.1%; Score 78.4; DB 10; Length 682;
Best Local Similarity 93.2%; Pred. No. 3.1e-18;
Matches 82; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCCTCACAGCCTTTTCTACAATGGCATTCAAT 60
DB 572 CGGAGCCGAGTCCTGTATCAGCCCTTTATCCTCACAGCCTTTTCTACAATGGCATTCAAT 631
QY 61 AAAGTCACGCTGTTTCTTGGTACAAAA 88
DB 632 AAAGTCACGCTGTTTCTTGGAAATAAAA 659
RESULT 15
AAS95009
ID AAS95009 standard; DNA; 695 BP.
AC AAS95009;
XX
XX 14-FEB-2002 (first entry)
XX
XX Human DNA sequence #264 expressed during foam cell differentiation.
XX
XX Human, foam cell differentiation; atherosclerosis; cerebral stroke;
XX cardiovascular disorder; coronary artery disease; gene therapy; ds.
XX
XX Homo sapiens.
XX
XX WO200177389-A2.
XX
XX 18-OCT-2001.
XX
XX 04-APR-2001; 2001WO-US011128.
XX
```

```
PR 05-APR-2000; 2000US-0195106P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Shiffman D, Somogyi R, Lawn R, Seilhamer JJ, Porter GJ, Mikita T;
PI Tai J;
XX
XX WPI; 2002-010925/01.
DR
XX
XX Composition useful for diagnosis of conditions, disorders or diseases
PT associated with atherosclerosis, comprises several polynucleotides that
PT are differentially expressed in foam cell development.
XX
XX Claim 1; Page 303-304; 315pp; English.
PS
XX
CC The present invention relates to the isolation of human polynucleotide
CC sequences that are differentially expressed during foam cell
CC differentiation. The polynucleotide sequences of the invention or a
CC composition comprising these polynucleotides are useful as a high
CC throughput method for detecting altered expression of one or more
CC polynucleotides in a sample. The polynucleotides can be used in the
CC diagnosis of disorders associated with foam cell development such as
CC atherosclerosis, cerebral stroke, and cardiovascular disorders such as
CC coronary artery disease. The polynucleotide sequences can also be used as
CC PCR primers and probes. The polynucleotides of the invention are also
CC useful in gene therapy. AAS94746-AAS95021 represent the human
CC polynucleotide sequences of the invention which are differentially
CC expressed during foam cell differentiation
XX
XX Sequence 695 BP; 158 A; 226 C; 166 G; 145 T; 0 U; 0 Other;
SQ
Query Match 89.1%; Score 78.4; DB 6; Length 695;
Best Local Similarity 93.2%; Pred. No. 3.1e-18;
Matches 82; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCCTCACAGCCTTTTCTACAATGGCATTCAAT 60
DB 585 CGGAGCCGAGTCCTGTATCAGCCCTTTATCCTCACAGCCTTTTCTACAATGGCATTCAAT 644
QY 61 AAAGTCACGCTGTTTCTTGGTACAAAA 88
DB 645 AAAGTCACGCTGTTTCTTGGAAATAAAA 672
Search completed: December 7, 2005, 00:03:04
Job time : 141.343 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 11:31:40 ; Search time 219.122 Seconds  
(without alignments)  
3320.997 Million cell updates/sec

Title: US-09-980-046B-5  
Perfect score: 88  
Sequence: 1 cggagtcgagtcctgtatca.....cgtgttcttgggtacaaaaa 88

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_NA\_Main:  
1: /cgn2\_6/ptodata/1/pubna/US07\_PUBCOMB.seq:  
2: /cgn2\_6/ptodata/1/pubna/US08\_PUBCOMB.seq:  
3: /cgn2\_6/ptodata/1/pubna/US09A\_PUBCOMB.seq:  
4: /cgn2\_6/ptodata/1/pubna/US09B\_PUBCOMB.seq:  
5: /cgn2\_6/ptodata/1/pubna/US10A\_PUBCOMB.seq:  
6: /cgn2\_6/ptodata/1/pubna/US10B\_PUBCOMB.seq:  
7: /cgn2\_6/ptodata/1/pubna/US10C\_PUBCOMB.seq:  
8: /cgn2\_6/ptodata/1/pubna/US10D\_PUBCOMB.seq:  
9: /cgn2\_6/ptodata/1/pubna/US10E\_PUBCOMB.seq:  
10: /cgn2\_6/ptodata/1/pubna/US11\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	ID	Description
1	80	90.9	483	3 US-09-918-995-21124
2	80	90.9	561	9 US-10-505-680-426
3	80	90.9	623	3 US-10-450-763-11581
4	80	90.9	642	3 US-09-738-973-119
5	80	90.9	642	3 US-09-854-133-119
6	80	90.9	642	5 US-10-144-649A-119
7	80	90.9	993	5 US-10-198-846-9827
8	78.4	89.1	682	5 US-10-084-817-16
9	78.4	89.1	682	5 US-10-240-965-264
10	75.8	86.1	118	7 US-10-242-535A-15601
11	75.8	86.1	118	7 US-10-085-783A-15601
12	75.8	86.1	118	8 US-10-661-242-3
13	75.8	86.1	596	6 US-10-115-831-9
14	75.8	86.1	808	5 US-10-197-666A-73
15	75.8	86.1	808	6 US-10-133-937-55
16	75.8	86.1	808	6 US-10-159-563-55
17	75.8	86.1	808	7 US-10-641-643-1279
18	75.8	86.1	808	9 US-10-929-182-31
19	75.8	86.1	808	9 US-10-505-680-426
20	75.8	86.1	808	9 US-10-956-157-2340
21	75.8	86.1	808	9 US-10-956-157-7575
22	75.4	85.7	401	7 US-10-242-535A-40680
23	75.4	85.7	401	7 US-10-085-783A-40680

Sequence 71, Appl  
Sequence 75, Appl  
Sequence 427, App  
Sequence 6787, Ap  
Sequence 6787, Ap  
Sequence 279728,  
Sequence 279728,  
Sequence 279728,  
Sequence 105978,  
Sequence 105978,  
Sequence 18813, A  
Sequence 18814, A  
Sequence 18815, A  
Sequence 18816, A  
Sequence 18813, A  
Sequence 18814, A  
Sequence 18815, A  
Sequence 18816, A  
Sequence 41318, A  
Sequence 41318, A  
Sequence 11580, A

24 75.4 85.7 606 5 US-10-197-666A-71  
25 75.4 85.7 621 5 US-10-197-666A-75  
26 74.8 85.0 545 9 US-10-505-680-427  
27 72.4 82.3 346 7 US-10-242-535A-6787  
28 72.4 82.3 346 7 US-10-085-783A-6787  
29 72 81.8 546 5 US-10-027-632-279728  
30 72 81.8 546 5 US-10-027-632-279729  
31 72 81.8 546 6 US-10-027-632-279728  
32 72 81.8 546 6 US-10-027-632-279729  
33 72 81.8 575 5 US-10-027-632-105978  
34 72 81.8 575 6 US-10-027-632-105978  
35 72 81.8 714 5 US-10-027-632-19813  
36 72 81.8 714 5 US-10-027-632-19814  
37 72 81.8 714 5 US-10-027-632-19815  
38 72 81.8 714 5 US-10-027-632-19816  
39 72 81.8 714 6 US-10-027-632-19813  
40 72 81.8 714 6 US-10-027-632-19814  
41 72 81.8 714 6 US-10-027-632-19815  
42 72 81.8 714 6 US-10-027-632-19816  
43 71.8 81.6 189 7 US-10-242-535A-41318  
44 71.8 81.6 189 7 US-10-085-783A-41318  
45 70.6 80.2 627 9 US-10-450-763-11580

## ALIGNMENTS

RESULT 1  
US-09-918-995-21124  
; Sequence 21124, Application US/09918995  
; Publication No. US20030073623A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc.  
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED  
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES  
; FILE REFERENCE: 20411-756  
; CURRENT APPLICATION NUMBER: US/09/918,995  
; PRIOR FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: US/09/235,076  
; PRIOR FILING DATE: 1999-01-20  
; NUMBER OF SEQ ID NOS: 38054  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 21124  
; LENGTH: 483  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)...(483)  
; OTHER INFORMATION: n = A,T,C or G

Query Match 90.9%; Score 80; DB 3; Length 483;  
Best Local Similarity 94.3%; Pred. No. 2.2e-18;  
Matches 83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
Qy 1 CGGAGTCGAGTCCTGATCAGCCCTTTCTCTACAGCTTTCTACAAATGCATTCAAT 60  
Db 217 CGGAGCGAGTCCTGATCAGCCCTTTCTCTACAGCTTTCTACAAATGCATTCAAT 276  
Qy 61 AAAGTGACGCTGTTTCTTGGTACAAAAA 88  
Db 277 AAAGTGACGCTGTTTCTTGGTACAAAAA 304

RESULT 2  
US-10-505-680-426/c  
; Sequence 426, Application US/10505680  
; Publication No. US20050095592A1  
; GENERAL INFORMATION:  
; APPLICANT: The Government of the United States of America as represented by the  
; APPLICANT: Secretary of the Department of Health and Human Services  
; APPLICANT: Jazaeri, Amir A.

```
; APPLICANT: Boyd, Jeff
; APPLICANT: Liu, Edison T.
; TITLE OF INVENTION: IDENTIFICATION OF OVARIAN CANCER TUMOR MARKERS AND THERAPEUTIC TA
; FILE REFERENCE: 4239-64452
; CURRENT APPLICATION NUMBER: US/10/505,680
; CURRENT FILING DATE: 2004-08-12
; PRIOR APPLICATION NUMBER: 60/357,031
; PRIOR FILING DATE: 2002-02-13
; NUMBER OF SEQ ID NOS: 822
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 426
; LENGTH: 561
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-505-680-426

Query Match          90.9%; Score 80; DB 9; Length 561;
Best Local Similarity 94.3%; Pred. No. 2.3e-18;
Matches 83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCCTCACAGCCTTTTCTACAATGGCATTCAAT 60
Db 88 CGGAGCGGAGTCCTGTATCAGCCCTTTATCCTCACAGCCTTTTCTACAATGGCATTCAAT 29

QY 61 AAAGTGCACGTGTTCTTGTGTAACAAAA 88
Db 28 AAAGTGCACGTGTTCTGTGTAACAAAAA 1

RESULT 3
US-10-450-763-11581
; Sequence 11581, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 11581
; LENGTH: 623
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURES:
; NAME/KEY: SIMILAR
; LOCATION: (49)..(366)
; OTHER INFORMATION: 96% homologous to Homo sapiens 1-8U, accession number
; OTHER INFORMATION: X57352, Smith-Waterman Score=541.
US-10-450-763-11581

Query Match          90.9%; Score 80; DB 9; Length 623;
Best Local Similarity 94.3%; Pred. No. 2.4e-18;
Matches 83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCCTCACAGCCTTTTCTACAATGGCATTCAAT 60
Db 535 CGGAGCGGAGTCCTGTATCAGCCCTTTATCCTCACAGCCTTTTCTACAATGGCATTCAAT 594

QY 61 AAAGTGCACGTGTTCTTGTGTAACAAAA 88
Db 595 AAAGTGCACGTGTTCTGTGTAACAAAAA 622

RESULT 4
US-09-738-973-119
; Sequence 119, Application US/09738973
; Patent No. US20020110563A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Fling, Steven P.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Algate, Paul A.
; APPLICANT: Secrist, Heather
; APPLICANT: Indrias, Carol Yoseph
; APPLICANT: Benson, Darin R.
; APPLICANT: Elliot, Mark
; APPLICANT: Mannion, Jane
; APPLICANT: Kalos, Michael D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C9
; CURRENT APPLICATION NUMBER: US/09/738,973
; CURRENT FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 587
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 119
; LENGTH: 642
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-738-973-119

Query Match          90.9%; Score 80; DB 3; Length 642;
Best Local Similarity 94.3%; Pred. No. 2.4e-18;
Matches 83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCCTCACAGCCTTTTCTACAATGGCATTCAAT 60
Db 535 CGGAGCGGAGTCCTGTATCAGCCCTTTATCCTCACAGCCTTTTCTACAATGGCATTCAAT 594

QY 61 AAAGTGCACGTGTTCTTGTGTAACAAAA 88
Db 595 AAAGTGCACGTGTTCTGTGTAACAAAAA 622

RESULT 5
US-09-854-133-119
; Sequence 119, Application US/09854133
; Publication No. US20020183499A1
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Henderson, Robert A.
; APPLICANT: Benson, Darin R.
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C10
; CURRENT APPLICATION NUMBER: US/09/854,133
; CURRENT FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 735
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 119
; LENGTH: 642
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-854-133-119

Query Match          90.9%; Score 80; DB 3; Length 642;
Best Local Similarity 94.3%; Pred. No. 2.4e-18;
Matches 83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCCTCACAGCCTTTTCTACAATGGCATTCAAT 60
Db 535 CGGAGCGGAGTCCTGTATCAGCCCTTTATCCTCACAGCCTTTTCTACAATGGCATTCAAT 594

QY 61 AAAGTGCACGTGTTCTTGTGTAACAAAA 88
```



```
Db 595 AAAGTGCACGTGTTCTTCTGGTGAAAAAA 622
|||||
RESULT 6
US-10-144-649A-119
; Sequence 119, Application US/10144649A
; Publication No. US20030118599A1
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Algate, Paul A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C11
; CURRENT APPLICATION NUMBER: US/10/144,649A
; CURRENT FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 749
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 119
; LENGTH: 642
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-144-649A-119
Query Match 90.9%; Score 80; DB 5; Length 642;
Best Local Similarity 94.3%; Pred. No. 2.4e-18;
Matches 83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCCTCACACGCTTTTCTACAAATGGCATTCAAT 60
Db 535 CGGAGCCGAGTCCTGTATCAGCCCTTTATCCTCACACGCTTTTCTACAAATGGCATTCAAT 594

Qy 61 AAAGTGCACGTGTTCTTCTGGTGACAAAAA 88
|||||
Db 595 AAAGTGCACGTGTTCTTCTGGTGAAAAAA 622

RESULT 7
US-10-198-846-9827
; Sequence 9827, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9827
; LENGTH: 993
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-198-846-9827
Query Match 90.9%; Score 80; DB 5; Length 993;
Best Local Similarity 94.3%; Pred. No. 2.8e-18;
Matches 83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCCTCACACGCTTTTCTACAAATGGCATTCAAT 60
Db 488 CGGAGCCGAGTCCTGTATCAGCCCTTTATCCTCACACGCTTTTCTACAAATGGCATTCAAT 547

Qy 61 AAAGTGCACGTGTTCTTCTGGTGACAAAAA 88
|||||
Db 548 AAAGTGCACGTGTTCTTCTGGTGAAAAAA 575

RESULT 8
US-10-084-817-16
; Sequence 16, Application US/10084817
; Publication No. US20030119009A1
; GENERAL INFORMATION:
; APPLICANT: Susan Stuart
; APPLICANT: Jed G. Nuchtern
; APPLICANT: Sharon E. Plon
; APPLICANT: Jason M. Shohet
; TITLE OF INVENTION: GENES REGULATED BY MYCN ACTIVATION
; FILE REFERENCE: PA-0046 US
; CURRENT APPLICATION NUMBER: US/10/084,817
; CURRENT FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 60/270,784
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 365
; SOFTWARE: PERL Program
; SEQ ID NO 16
; LENGTH: 682
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030119009A1 1101453.2
US-10-084-817-16
Query Match 89.1%; Score 78.4; DB 5; Length 682;
Best Local Similarity 93.2%; Pred. No. 9.4e-18;
Matches 82; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCCTCACACGCTTTTCTACAAATGGCATTCAAT 60
Db 572 CGGAGCCGAGTCCTGTATCAGCCCTTTATCCTCACACGCTTTTCTACAAATGGCATTCAAT 631

Qy 61 AAAGTGCACGTGTTCTTCTGGTGACAAAAA 88
|||||
Db 632 AAAGTGCACGTGTTCTTCTGGTGAAAAA 659

RESULT 9
US-10-240-965-264
; Sequence 264, Application US/10240965
; Publication No. US20030165924A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: SHIFFMAN, Dov
; APPLICANT: SOMOGYI, Roland
; APPLICANT: LAWN, Richard M.
; APPLICANT: SEILHAWER, Jeffrey J.
; APPLICANT: PORTER, Gordon J.
; APPLICANT: MIKITA, Thomas
; APPLICANT: TAI, Julie
; TITLE OF INVENTION: GENES EXPRESSED IN FOAM CELL DIFFERENTIATION
; FILE REFERENCE: PA-0025 PCT
; CURRENT APPLICATION NUMBER: US/10/240,965
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: 60/195,106
; PRIOR FILING DATE: 2000-04-05
; NUMBER OF SEQ ID NOS: 276
; SOFTWARE: PERL Program
; SEQ ID NO 264
; LENGTH: 695
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030165924A1 444902.6
US-10-240-965-264
```



; SOFTWARE: pt\_FL\_genes Version 2.0  
; SEQ ID NO 9  
; LENGTH: 596  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (22)..(423)  
US-10-115-831-9

Query Match 86.1%; Score 75.8; DB 6; Length 596;  
Best Local Similarity 97.5%; Pred. No. 7.7e-17;  
Matches 77; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCTCTCACACGCTTTTCTACAAATGGCATTCAAT 60  
|||||  
Db 506 CGGAGCCGAGTCCTGTATCAGCCCTTTATCTCTCACACGCTTTTCTACAAATGGCATTCAAT 565  
|||||  
Qy 61 AAAGTGCACGTGTTTCTTG 79  
|||||  
Db 566 AAAGTGCACGTGTTTCTGG 584  
|||||

RESULT 14  
US-10-197-666A-73  
; Sequence 73, Application US/10197666A  
; Publication No. US20030092037A1  
; GENERAL INFORMATION:  
; APPLICANT: ASAH KASEI KABUSIKI KAISYA  
; TITLE OF INVENTION: E1k1 phosphorylation related gene  
; FILE REFERENCE: PH-1548US  
; CURRENT APPLICATION NUMBER: US/10/197,666A  
; CURRENT FILING DATE: 2002-11-18  
; PRIOR APPLICATION NUMBER: JP 2001-218204  
; PRIOR FILING DATE: 2001-07-18  
; PRIOR APPLICATION NUMBER: JP 2001-263450  
; PRIOR FILING DATE: 2001-08-31  
; PRIOR APPLICATION NUMBER: JP 2002-012176  
; PRIOR FILING DATE: 2002-01-21  
; PRIOR APPLICATION NUMBER: US 60/305,884  
; PRIOR FILING DATE: 2001-07-18  
; PRIOR APPLICATION NUMBER: US 60/316,304  
; PRIOR FILING DATE: 2001-09-04  
; PRIOR APPLICATION NUMBER: US 60/350,027  
; PRIOR FILING DATE: 2002-01-23  
; NUMBER OF SEQ ID NOS: 156  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 73  
; LENGTH: 808  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (238)..(636)  
US-10-197-666A-73

Query Match 86.1%; Score 75.8; DB 5; Length 808;  
Best Local Similarity 97.5%; Pred. No. 8.7e-17;  
Matches 77; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCTCTCACACGCTTTTCTACAAATGGCATTCAAT 60  
|||||  
Db 722 CGGAGCCGAGTCCTGTATCAGCCCTTTATCTCTCACACGCTTTTCTACAAATGGCATTCAAT 781  
|||||  
Qy 61 AAAGTGCACGTGTTTCTTG 79  
|||||  
Db 782 AAAGTGCACGTGTTTCTGG 800  
|||||

RESULT 15  
US-10-133-937-55  
; Sequence 55, Application US/10133937  
; Publication No. US20030207278A1

; GENERAL INFORMATION:  
; APPLICANT: Khan, Javed  
; APPLICANT: Ringner, Markus  
; APPLICANT: Peterson, Carsten  
; APPLICANT: Meltzer, Paul  
; TITLE OF INVENTION: METHODS FOR ANALYZING HIGH DIMENSIONAL DATA FOR CLASSIFYING,  
; TITLE OF INVENTION: DIAGNOSING, PROGNOSTICATING, AND/OR PREDICTING DISEASES AND  
; TITLE OF INVENTION: OTHER BIOLOGICAL STATES  
; FILE REFERENCE: 11613.56US01  
; CURRENT APPLICATION NUMBER: US/10/133,937  
; CURRENT FILING DATE: 2002-11-04  
; NUMBER OF SEQ ID NOS: 99  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 55  
; LENGTH: 808  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-133-937-55

Query Match 86.1%; Score 75.8; DB 6; Length 808;  
Best Local Similarity 97.5%; Pred. No. 8.7e-17;  
Matches 77; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCTCTCACACGCTTTTCTACAAATGGCATTCAAT 60  
|||||  
Db 722 CGGAGCCGAGTCCTGTATCAGCCCTTTATCTCTCACACGCTTTTCTACAAATGGCATTCAAT 781  
|||||  
Qy 61 AAAGTGCACGTGTTTCTTG 79  
|||||  
Db 782 AAAGTGCACGTGTTTCTGG 800  
|||||

Search completed: December 6, 2005, 22:57:05  
Job time : 220.122 secs

***This Page Blank (Lup16)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 11:34:11 ; Search time 60.0707 Seconds  
(without alignments)  
455.919 Million cell updates/sec

Title: US-09-980-046B-5  
Perfect score: 88  
Sequence: 1 cggagtcgagtcctgtatca.....cgtgttcttggtacaaaaa 88

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3289935 seqs, 155610033 residues

Total number of hits satisfying chosen parameters: 6579870

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA New:  
1: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:  
2: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:  
3: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:  
4: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:  
5: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:  
6: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:  
7: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:  
8: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:  
9: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:  
10: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	41.2	46.8	182314	7	US-11-112-908-45
C 2	30	34.1	4116	6	US-10-750-185-38038
C 3	27.2	30.9	1033	6	US-10-750-185-57623
C 4	25.6	29.1	1668	6	US-10-750-185-50602
C 5	25.4	28.9	2224	6	US-10-750-185-53492
C 6	25.2	28.6	1642	6	US-10-750-185-41665
C 7	24.8	28.2	1998	6	US-10-750-185-45955
C 8	24.6	28.0	494	6	US-10-750-185-2846
C 9	24.2	27.5	2156	6	US-10-750-185-51326
C 10	24.2	27.5	2357	6	US-10-750-185-56680
C 11	24	27.3	3541	6	US-10-750-185-24928
C 12	23.8	27.0	1049	6	US-10-750-185-64427
C 13	23.8	27.0	2248	6	US-10-750-185-25851
C 14	23.8	27.0	2301	6	US-10-793-626-419
C 15	23.8	27.0	2744	6	US-10-793-626-4069
C 16	23.8	27.0	3127	6	US-10-793-626-3502
C 17	23.8	27.0	3170	6	US-10-793-626-3892
C 18	23.8	27.0	4094	6	US-10-793-626-3641
C 19	23.8	27.0	6968	6	US-10-793-626-759
C 20	23.6	26.8	1759	6	US-10-750-185-57531
C 21	23.4	26.6	1407	6	US-10-750-185-34679
C 22	23.4	26.6	1506	6	US-10-750-185-44563
C 23	23.4	26.6	1987	6	US-10-750-185-62265

ALIGNMENTS

RESULT 1

US-11-112-908-45/c  
; Sequence 45, Application US/11112908  
; Publication No. US20050260659A1  
; GENERAL INFORMATION:  
; APPLICANT: Harris, Cole  
; TITLE OF INVENTION: Breast Cancer Biomarkers  
; FILE REFERENCE: 04-164-US  
; CURRENT APPLICATION NUMBER: US/11/112,908  
; CURRENT FILING DATE: 2005-04-22  
; PRIOR APPLICATION NUMBER: US 60/564,758  
; PRIOR FILING DATE: 2004-04-23  
; PRIOR APPLICATION NUMBER: US 60/575,978  
; PRIOR FILING DATE: 2004-06-01  
; PRIOR APPLICATION NUMBER: US 60/631,702  
; PRIOR FILING DATE: 2004-11-30  
; PRIOR APPLICATION NUMBER: US 60/633,826  
; PRIOR FILING DATE: 2004-12-07  
; NUMBER OF SEQ ID NOS: 511  
; SOFTWARE: Patentin version 3.3  
; SEQ ID NO 45  
; LENGTH: 182314  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-11-112-908-45

Query Match 46.8%; Score 41.2; DB 7; Length 182314;  
Best Local Similarity 74.3%; Pred. No. 3.2e-05;  
Matches 52; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 19 CAGCCCTTTTCTCCACACGCTTTTCTACATGCGATCAATGAAGTCGCGTCTTCTT 78

Db 27145 CAGCAGTTTACTACATACATCTTTTCTACATGTTTGTGATGTCGCGTCTTCTT 27086

Qy 79 GGTACAAAAA 88

Db 27085 AAAAAAAA 27076

RESULT 2

US-10-750-185-38038  
; Sequence 38038, Application US/10750185  
; Publication No. US20050260603A1  
; GENERAL INFORMATION:  
; APPLICANT: MMI GENOMICS, INC.

```
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 38038
; LENGTH: 4116
; TYPE: DNA
; ORGANISM: Bovine 19866881138467
US-10-750-185-38038

Query Match      34.1%; Score 30; DB 6; Length 4116;
Best Local Similarity 72.2%; Pred. No. 0.11;
Matches 39; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 22 CCCTTTATCTCTCACACGCTTTTCTACAAATGGCATTCAATAAAGTGACGTTT 75
Db 1469 CCGTCTATCTGATATGCTTTCTCCCAATGGCTTTCTACAGTGAACAGTTT 1522

RESULT 3
US-10-750-185-57623
; Sequence 57623, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 57623
; LENGTH: 1033
; TYPE: DNA
; ORGANISM: Bovine 19866881623743
US-10-750-185-57623

Query Match      30.9%; Score 27.2; DB 6; Length 1033;
Best Local Similarity 64.1%; Pred. No. 0.73;
Matches 41; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Qy 22 CCCTTTATCTCTCACACGCTTTTCTACAAATGGCATTCAATAAAGTGACGTTTCTTTGGT 81
Db 177 CCCTTTAACACACCTCTCTTTTAACTCTGTCATTCTTCTGTCATATTTTCAAGAT 236

Qy 82 ACAA 85
Db 237 CCTA 240

RESULT 4
US-10-750-185-50602
; Sequence 50602, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
```

```
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 50602
; LENGTH: 1668
; TYPE: DNA
; ORGANISM: Bovine 19866880988770
US-10-750-185-50602

Query Match      29.1%; Score 25.6; DB 6; Length 1668;
Best Local Similarity 70.8%; Pred. No. 3.3;
Matches 34; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 41 TTTCTCAATGGCATTCAATAAAGTGACGTTTCTTTGGTACAAAAA 88
Db 222 TTTCTCAATGACATTCAATAAAGTCCAGGCTGGAGGNAAAAAA 269

RESULT 5
US-10-750-185-53492/c
; Sequence 53492, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 53492
; LENGTH: 2224
; TYPE: DNA
; ORGANISM: Bovine 19866880905627
US-10-750-185-53492

Query Match      28.9%; Score 25.4; DB 6; Length 2224;
Best Local Similarity 64.4%; Pred. No. 4.3;
Matches 38; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

Qy 11 TCCTGTATCAGCCCTTTATCCTCACACGCTTTTCTCAATGGCATTCAATAAAGTGAC 69
Db 983 TCCTTTATCAGCTTTTATATATATGATGATTTTAAAAAATAGCAATCCATAGAAAGAC 925

RESULT 6
US-10-750-185-41665/c
; Sequence 41665, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
```

; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MM1100-2  
; CURRENT APPLICATION NUMBER: US/10/750,185  
; CURRENT FILING DATE: 2003-12-31  
; PRIOR APPLICATION NUMBER: US 60/437,482  
; PRIOR FILING DATE: 2002-12-31  
; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 41665  
; LENGTH: 1642  
; TYPE: DNA  
; ORGANISM: Bovine 19866880829778  
US-10-750-185-41665

Query Match 28.6%; Score 25.2; DB 6; Length 1642;  
Best Local Similarity 62.9%; Pred. No. 4.6;  
Matches 39; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Qy 27 TATCCTACAGCCTTTTCTACAAATGCGATTCAATAAGTGCACGTGTTCTTGTGTCAAA 86  
Db 137 TTTTCTATCCCTCTTTTCTACTATAGCATTTCTCTCAAAAGCATTTCTTTCTTTCAATAA 78  
Qy 87 AA 88  
Db 77 AA 76

RESULT 7  
US-10-750-185-45955/c  
; Sequence 45955, Application US/10750185  
; Publication No. US20050260603A1  
; GENERAL INFORMATION:  
; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MM1100-2  
; CURRENT APPLICATION NUMBER: US/10/750,185  
; CURRENT FILING DATE: 2003-12-31  
; PRIOR APPLICATION NUMBER: US 60/437,482  
; PRIOR FILING DATE: 2002-12-31  
; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 45955  
; LENGTH: 1998  
; TYPE: DNA  
; ORGANISM: Bovine 19866880580954  
US-10-750-185-45955

Query Match 28.2%; Score 24.8; DB 6; Length 1998;  
Best Local Similarity 56.0%; Pred. No. 6.9;  
Matches 47; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

Qy 5 GTGAGTCCTGTATCAGCCCTTTATCCTCACACGCTTTTCTACATGCGATTGCAATAAG 64  
Db 1263 GTAGCTTCGTGAGCAGCAGGTTCATCTCTGCGCTTCTTAAAGCAATGCAAGAGC 1204  
Qy 65 TGCAGCTGTTTCTTGGTACAAAA 88  
Db 1203 TGTAGACTCTTAGCTATATAA 1180

RESULT 8  
US-10-750-185-2846  
; Sequence 2846, Application US/10750185

; Publication No. US20050260603A1  
; GENERAL INFORMATION:  
; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MM1100-2  
; CURRENT APPLICATION NUMBER: US/10/750,185  
; CURRENT FILING DATE: 2003-12-31  
; PRIOR APPLICATION NUMBER: US 60/437,482  
; PRIOR FILING DATE: 2002-12-31  
; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2846  
; LENGTH: 494  
; TYPE: DNA  
; ORGANISM: Bovine MMBT14110  
US-10-750-185-2846

Query Match 28.0%; Score 24.6; DB 6; Length 494;  
Best Local Similarity 57.0%; Pred. No. 5.1;  
Matches 45; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

Qy 9 AGTCTGTATCAGCCCTTTATCCTCACACGCTTTTCTACATGCGATTGCAATAAGTGA 68  
Db 178 ATTGCTGTAGGACCTCTCACTCTCTCACGCTTAATAGATGCAATAAAACCAAAAGA 237  
Qy 69 CGTGTCTTCTTGGTACAAAA 87  
Db 238 TTTTCTTTTCTCTGCTTAA 256

RESULT 9  
US-10-750-185-51326/c  
; Sequence 51326, Application US/10750185  
; Publication No. US20050260603A1  
; GENERAL INFORMATION:  
; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MM1100-2  
; CURRENT APPLICATION NUMBER: US/10/750,185  
; CURRENT FILING DATE: 2003-12-31  
; PRIOR APPLICATION NUMBER: US 60/437,482  
; PRIOR FILING DATE: 2002-12-31  
; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 51326  
; LENGTH: 2156  
; TYPE: DNA  
; ORGANISM: Bovine 19866880417584  
US-10-750-185-51326

Query Match 27.5%; Score 24.2; DB 6; Length 2156;  
Best Local Similarity 71.1%; Pred. No. 12;  
Matches 32; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 33 CACAGCTTTTCTACAATGCGATTCAATAAGTGCACGTGTTCT 77  
Db 1973 CACATTTCTTCCATACAGGCTTTCTGTTAAGTGCAGTGAATCT 1929

RESULT 10  
US-10-750-185-56680

```

; Sequence 56680, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 56680
; LENGTH: 2357
; TYPE: DNA
; ORGANISM: Bovine 19866880352383
US-10-750-185-56680

Query Match 27.5%; Score 24.2; DB 6; Length 2357;
Best Local Similarity 59.4%; Pred. No. 12;
Matches 41; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 20 AGCCCTTTATCCTCACACGCTTTTCTACAATGGCAATCAATAAGTGCACGTGTTCTTG 79
Db 1326 AGGCCTGGAATCAGGAGCGCTGGCACACACGGGCACTCAAGNAAGCTCTTGTTGAAG 1385

Qy 80 GTACAAAAA 88
Db 1386 GAACAACA 1394

RESULT 11
US-10-750-185-24928
; Sequence 24928, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 24928
; LENGTH: 3541
; TYPE: DNA
; ORGANISM: Bovine 19866880691281
US-10-750-185-24928

Query Match 27.3%; Score 24; DB 6; Length 3541;
Best Local Similarity 64.3%; Pred. No. 17;
Matches 36; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

Qy 11 TCCGTATCAGCCCTTTATCCTCACACGCTTTTCTACAATGGCAATCAATAAGTGCACGTG 66
Db 2645 TCCAGTTCAGCCCTATTTGCTCAGATGGTTCTCTAAAACCTTGAATTAATTCAGTG 2700

RESULT 12
US-10-750-185-64427
; Sequence 64427, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 64427
; LENGTH: 1049
; TYPE: DNA
; ORGANISM: Bovine 19866880882500
US-10-750-185-64427

Query Match 27.0%; Score 23.8; DB 6; Length 1049;
Best Local Similarity 57.3%; Pred. No. 13;
Matches 43; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

Qy 14 TGATCAGCCCTTTATCCTCACACGCTTTTCTACAATGGCAATCAATAAGTGCACGTGT 73
Db 805 TGAATTTATCCTAAATTTTAAAGATGTTTAAAGATGACATACATTAATGCTCATATGT 864

Qy 74 TTCTTGGTACAAAA 88
Db 865 GTTAGTTACAAAA 879

RESULT 13
US-10-750-185-25851/c
; Sequence 25851, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 25851
; LENGTH: 2248
; TYPE: DNA
; ORGANISM: Bovine 19866880712906
US-10-750-185-25851

Query Match 27.0%; Score 23.8; DB 6; Length 2248;
Best Local Similarity 59.7%; Pred. No. 17;
Matches 40; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

Qy 14 TGATCAGCCCTTTATCCTCACACGCTTTTCTACAATGGCAATCAATAAGTGCACGTGT 73
Db 1517 TATTTGATCCTTTTATTTTCAATGTTATATATATCTTGGTCTATAAAAAGACTGTTT 1458

Qy 74 TTCTTGG 80

```



Db 1457 TTCTGG 1451  
|||||||

RESULT 14  
US-10-793-626-419  
; Sequence 419, Application US/10793626  
; Publication No. US20050255478A1  
; GENERAL INFORMATION:  
; APPLICANT: KIMMERLY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
; FILE REFERENCE: PU3480US  
; CURRENT APPLICATION NUMBER: US/10/793.626  
; CURRENT FILING DATE: 2004-03-04  
; PRIOR APPLICATION NUMBER: 60/164,258  
; PRIOR FILING DATE: 1999-11-09  
; NUMBER OF SEQ ID NOS: 4472  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 419  
; LENGTH: 2301  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
; OTHER INFORMATION: nucleic acid sequence  
US-10-793-626-419

Query Match 27.0%; Score 23.8; DB 6; Length 2301;  
Best Local Similarity 72.1%; Pred. No. 17;  
Matches 31; Conservative 0; Mismatches 12; Indels 0; Gaps 0;  
  
QY 31 CTCACACGCTTTTCTACAAATGGCATTCAATAAAGTGACGTTGT 73  
|||||  
Db 1817 CTCAAACATTCATACAAATGGCATTCAAAAAGGGGAACGTTGT 1859  
|||||

RESULT 15  
US-10-793-626-4069/c  
; Sequence 4069, Application US/10793626  
; Publication No. US20050255478A1  
; GENERAL INFORMATION:  
; APPLICANT: KIMMERLY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
; FILE REFERENCE: PU3480US  
; CURRENT APPLICATION NUMBER: US/10/793.626  
; CURRENT FILING DATE: 2004-03-04  
; PRIOR APPLICATION NUMBER: 60/164,258  
; PRIOR FILING DATE: 1999-11-09  
; NUMBER OF SEQ ID NOS: 4472  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4069  
; LENGTH: 2744  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
; OTHER INFORMATION: nucleic acid sequence  
US-10-793-626-4069

Query Match 27.0%; Score 23.8; DB 6; Length 2744;  
Best Local Similarity 72.1%; Pred. No. 18;  
Matches 31; Conservative 0; Mismatches 12; Indels 0; Gaps 0;  
  
QY 31 CTCACACGCTTTTCTACAAATGGCATTCAATAAAGTGACGTTGT 73  
|||||  
Db 715 CTCAAACATTCATACAAATGGCATTCAAAAAGGGGAACGTTGT 673  
|||||

Search completed: December 6, 2005, 23:16:51  
Job time : 60.0707 secs

*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 21:43:57 ; Search time 1004.45 Seconds  
(without alignments)  
4099.008 Million cell updates/sec

Title: US-09-980-046B-5  
Perfect score: 88  
Sequence: 1 cggagtcgagtcctgtatca.....cgtgttcttctgtacaaaaa 88

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

- 1: gb\_est1:\*
- 2: gb\_est2:\*
- 3: gb\_est3:\*
- 4: gb\_est4:\*
- 5: gb\_est5:\*
- 6: gb\_est6:\*
- 7: gb\_est7:\*
- 8: gb\_est8:\*
- 9: gb\_est9:\*
- 10: gb\_est10:\*
- 11: gb\_est11:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	84.8	95.4	661	AV702819	AV702819 UI-HF-ES0
2	83.8	95.2	639	CFI29716	CFI29716 UI-HF-ES0
3	82.2	93.4	640	BF685813	BF685813 602140496
4	81.6	92.7	593	BI517742	BI517742 603042174
5	81.6	92.7	594	BO689205	BO689205 AGENCOURT
6	81.6	92.7	632	BI599867	BI599867 603246815
7	81.6	92.7	652	BF339806	BF339806 602038812
8	81.6	92.7	914	BI762444	BI762444 603048836
9	80.6	91.6	661	3 BQ067642	BQ067642 AGENCOURT
10	80	90.9	236	CA431735	CA431735 UI-H-DF0-
11	80	90.9	243	AI547782	AI547782 UI-R-C3-s
12	80	90.9	251	CN310115	CN310115 170060000
13	80	90.9	271	AA542950	AA542950 ni907.s
14	80	90.9	280	CN483092	CN483092 hw28c11.y
15	80	90.9	313	BI060666	BI060666 PMO-UT010
16	80	90.9	316	BM924345	BM924345 AGENCOURT
17	80	90.9	319	AW516688	AW516688 xD71a03.x
18	80	90.9	330	AW950085	AW950085 EST162275
19	80	90.9	347	AI754997	AI754997 cr34a04.x
20	80	90.9	351	AA573395	AA573395 nm53b11.s
21	80	90.9	358	1 AA657893	AA657893 nu07h03.s
22	80	90.9	358	3 BM543768	BM543768 AGENCOURT

c	23	80	90.9	358	8	N29647	N29647 yw73b02.s1
	24	80	90.9	362	5	CFI29432	CFI29432 UI-HF-ES0
	25	80	90.9	363	5	EX118214	EX118214 BX118214
	26	80	90.9	387	6	CF272419	CF272419 AGENCOURT
	27	80	90.9	393	2	BI257614	BI257614 602967192
	28	80	90.9	393	5	BUI53040	BUI53040 AGENCOURT
	29	80	90.9	399	2	EG059171	EG059171 nah46R05.
c	30	80	90.9	404	1	AW023437	AW023437 df54h03.y
	31	80	90.9	405	1	AW472903	AW472903 x822a08.x
	32	80	90.9	406	3	BM850332	BM850332 K-EST0130
	33	80	90.9	410	1	AA658408	AA658408 nu18d02.s
	34	80	90.9	412	1	AI862050	AI862050 tw71b08.x
	35	80	90.9	423	1	AA534111	AA534111 nj73g01.s
	36	80	90.9	428	8	T55566	T55566 yb39c05.s1
	37	80	90.9	432	1	AA535668	AA535668 nj77g12.s
	38	80	90.9	437	3	BM672821	BM672821 UI-E-CQ0-
	39	80	90.9	439	1	AL036296	AL036296 DKF2p564N
	40	80	90.9	442	3	BM741599	BM741599 K-EST0014
	41	80	90.9	446	5	BU728111	BU728111 UI-E-CQ0-
	42	80	90.9	449	1	AA535290	AA535290 nj79b05.s
	43	80	90.9	450	8	N93952	N93952 zb74604.s1
	44	80	90.9	458	3	BQ029595	BQ029595 UI-H-DF0-
	45	80	90.9	458	5	BU737184	BU737184 UI-E-DX1-

ALIGNMENTS

RESULT 1  
AV702819  
LOCUS  
DEFINITION AV702819 ADB Homo sapiens cDNA clone ADBDGH07 5', mRNA linear EST 08-OCT-2000  
ACCESSION AV702819  
VERSION AV702819.1 GI:10719149  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE 1 (bases 1 to 661)  
AUTHORS Peng, Y., Song, H., Huang, Q., Huang, C., Gu, Y., Yang, Y., Gao, G., Xiao, H., Xu, X., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z., Zeng, L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G., Hu, R., Chen, J., Chen, Z. and Han, Z.  
TITLE Homo sapiens cDNA ADB clones  
JOURNAL Unpublished (2000)  
COMMENT Contact: Zeguang Han  
Chinese National Human Genome Center at Shanghai  
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai  
201203, P. R. China  
Tel: 86-21-50801919(ex.45)  
Fax: 86-21-50801922  
Email: hanzgchgc.sh.cn  
This clone is available at CHGC in Shanghai.  
FEATURES  
Location/Qualifiers  
1..661  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="ADB07"  
/tissue type="Adrenal gland"  
/dev stage="Adult"  
/lab\_host="SOLR"  
/clone\_lib="ADB"  
/note="Vector: pBluescript sk(-); Site\_1: EcoRI; Site\_2: XhoI"

ORIGIN

Query Match 96.4%; Score 84.8; DB 1; Length 661;  
Best Local Similarity 97.7%; Pred. No. 7.5e-18;  
Matches 86; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 CGAGTCCGAGTCCTGTATCAGCCCTTTATCTCTCACAGCTTTTCTACAATGGCATTCAAT 60
Db 477 CGGAGCCGAGTCCTGTATCAGCCCTTTATCTCTCACAGCTTTTCTACAATGGCATTCAAT 536

QY 61 AAGTGCACGCTGTTCTTCTGGTACAAAAA 88
Db 537 AAGTGCACGCTGTTCTTCTGGTAAAAAAA 564

RESULT 2
CF129716 639 bp mRNA linear EST 05-AUG-2003
LOCUS UI-HF-ESO-avv-p-11-0-UI.r1 NIH_MGC_213 Homo sapiens cDNA clone
DEFINITION IMAGE:30565138 5', mRNA sequence.
ACCESSION CF129716
VERSION CF129716.1 GI:33210245
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 639)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
JOURNAL discovery
PUBMED Genome Res. 6 (9), 791-806 (1996)
COMMENT Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Mary Hendrix
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/humanf1.html
Seq primer: pYX-5.

FEATURES
source
1..639
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30565138"
/tissue_type="Chondrosarcoma Lung Metastasis cell lines"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_MGC_213"
/notes="Organ: Lung; Vector: pYX-Aac; Site 1: EcoR I;
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with Not I and then cloned
directionally into pYX-Aac vector. The library tag
sequence located between the Not I site and the polyA tail
is GATTAAGGCCA. Tissue was provided by Mary Hendrix."

ORIGIN
Query Match 95.2%; Score 83.8; DB 6; Length 639;
Best Local Similarity 96.6%; Pred. No. 1.6e-17;
Matches 85; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CGAGTCCGAGTCCTGTATCAGCCCTTTATCTCTCACAGCTTTTCTACAATGGCATTCAAT 60
Db 538 CGGAGCCGAGTCCTGTATCAGCCCTTTATCTCTCACAGCTTTTCTACAATGGCATTCAAT 597

QY 61 AAGTGCACGCTGTTCTTCTGGTACAAAAA 88

```

```

Db 598 AAGTGCACGCTGTTCTTCTGGTAAAAAAA 625

RESULT 3
BF685813 640 bp mRNA linear EST 22-DEC-2000
LOCUS 602140496F1 NIH_MGC_46 Homo sapiens cDNA clone IMAGE:4301807 5',
DEFINITION mRNA sequence.
ACCESSION BF685813
VERSION BF685813.1 GI:11971221
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 640)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strauberg, Ph.D.
Email: cga@rs-x@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LICM1162 row: 0 column: 24
High quality sequence stop: 638.
Location/Qualifiers
1..640
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4301807"
/tissue_type="Telomysarcoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_46"
/notes="Organ: uterus; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."

ORIGIN
Query Match 93.4%; Score 82.2; DB 2; Length 640;
Best Local Similarity 95.5%; Pred. No. 5.7e-17;
Matches 84; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGAGTCCGAGTCCTGTATCAGCCCTTTATCTCTCACAGCTTTTCTACAATGGCATTCAAT 60
Db 535 CGGAGCCGAGTCCTGTATCAGCCCTTTATCTCTCACAGCTTTTCTACAATGGCATTCAAT 594

QY 61 AAGTGCACGCTGTTCTTCTGGTACAAAAA 88
Db 595 AAGTGCACGCTGTTCTTCTGGTACAAAAA 622

RESULT 4
BI517742 593 bp mRNA linear EST 29-AUG-2001
LOCUS 603042174F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5182435 5',
DEFINITION mRNA sequence.
ACCESSION BI517742
VERSION BI517742
KEYWORDS EST.

```

```

SOURCE      Homo sapiens (human)
ORGANISM    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homnidae; Homo.
REFERENCE    1 (bases 1 to 593)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11455 row: d column: 20
            High quality sequence stop: 561.

FEATURES    Location/Qualifiers
            1..593
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:5182435"
            /lab_host="DH10B"
            /clone_lib="NIH_MGC_116"
            /note="Organ: pooled colon, kidney, stomach; Vector:
            pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
            source anonymous pool of 3 colons, age 26 yo male, 49 yo
            female, 71 yo male colon; 46 yo male kidney, and pool of 2
            stomachs, 62 yo male and 70 yo female. Library is
            oligo-dT primed and directionally cloned (EcoRV site is
            destroyed upon cloning). Average insert size 1.4 kb,
            insert size range 1-3 kb. Library is normalized and
            enriched for full-length clones and was constructed by C.
            Gruber (Invitrogen). Research Genetics tracking code
            023. Note: this is a NIH_MGC Library."

ORIGIN
Query Match      92.7%; Score 81.6; DB 3; Length 593;
Best Local Similarity 95.5%; Pred. No. 8.9e-17;
Matches 84; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGAGTCGAGTCCTGTATCAGCCCTTTATCTCTCACACGCTTTCTACAAATGGCATTCAAT 60
    |||||
DB 482 CGGAGCCGAGTCCTGTATCAGCCCTTTATCTCTCACACGCTTTCTACAAATGGCATTCAAT 541
    |||||

QY 61 AAAGTGCACGCTGTTCTTGTTGTAACAAAA 88
    |||||
DB 542 AAAGTGCACGCTGTTCTTGTTGTAACAAAA 569
    |||||

RESULT 5
LOCUS      BQ689205          594 bp mRNA linear EST 15-JUL-2002
DEFINITION AGENCOURT_8062029 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6210960
            5', mRNA sequence.
ACCESSION  BQ689205
VERSION     BQ689205.1 GI:21814521
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homnidae; Homo.
REFERENCE   1 (bases 1 to 594)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov

SOURCE      Homo sapiens (human)
ORGANISM    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homnidae; Homo.
REFERENCE    1 (bases 1 to 594)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2372 row: l column: 01
High quality sequence stop: 546.

FEATURES    Location/Qualifiers
            1..594
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clones="IMAGE:6210960"
            /tissue_type="ductal carcinoma, cell line"
            /lab_host="DH10B (phage-resistant)"
            /clone_lib="NIH_MGC_110"
            /note="Organ: pancreas; Vector: pOTB7; Site 1: XhoI;
            Site 2: EcoRI; cDNA made by oligo-dT priming.
            Directionally cloned into EcoRI/XhoI sites using the
            following 5' adaptor: GGCAGAG(G). Library constructed by
            Ling Hong in the laboratory of Gerald M. Rubin (University
            of California, Berkeley) using ZAP-cDNA synthesis kit
            (Stratagene) and Superscript II RT (Life Technologies).
            Note: this is a NIH_MGC Library."

ORIGIN
Query Match      92.7%; Score 81.6; DB 5; Length 594;
Best Local Similarity 95.5%; Pred. No. 9e-17;
Matches 84; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCTCTCACACGCTTTCTACAAATGGCATTCAAT 60
    |||||
DB 497 CGGAGCCGAGTCCTGTATCAGCCCTTTATCTCTCACACGCTTTCTACAAATGGCATTCAAT 556
    |||||

QY 61 AAAGTGCACGCTGTTCTTGTTGTAACAAAA 88
    |||||
DB 557 AAAGTGCACGCTGTTCTTGTTGTAACAAAA 584
    |||||

RESULT 6
LOCUS      B1599867          632 bp mRNA linear EST 07-SEP-2001
DEFINITION 603246815F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5294762 5',
            mRNA sequence.
ACCESSION  B1599867
VERSION     B1599867.1 GI:15492806
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homnidae; Homo.
REFERENCE   1 (bases 1 to 632)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
            cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11745 row: m column: 03
            High quality sequence stop: 632.

FEATURES    Location/Qualifiers
            1..632
            /organism="Homo sapiens"

```



Db 626 AAGTGCACGTGTTTCTGTCGCAAAA 653

RESULT 9  
BQ067642 661 bp mRNA linear EST 02-APR-2002  
LOCUS AGENCOURT\_6759133 NIH\_MGC\_115 Homo sapiens cDNA clone IMAGE:5755219  
DEFINITION 5', mRNA sequence.

ACCESSION BQ067642  
VERSION BQ067642.1 GI:19896688  
KEYWORDS EST.

SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo

REFERENCE 1 (bases 1 to 661)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs@mail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov  
Plate: LLAM12793 row: n column: 20  
High quality sequence stop: 632.

FEATURES  
Location/Qualifiers  
1..661  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5755219"  
/lab\_host="DH10B"  
/clone\_lib="NIH\_MGC\_115"  
/note="Organ: pooled brain, lung, testis; Vector: pCMV-SPORT6; Site\_1: NotI; Site\_2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27; 1 male lung, age 27; and 1 male testis, age 69. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH\_MGC Library."

ORIGIN  
Query Match 91.6%; Score 80.6; DB 3; Length 661;  
Best Local Similarity 94.3%; Pred. No. 2e-16;  
Matches 83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCTCTCACACGCTTTTCTACAAATGCAATCAAT 60  
|||||  
Db 540 CGGAGCCGAGTCCTGTATCAGCCCTTTATCTCTCACACGCTTTTCTACAAATGCAATCAAT 599  
|||||

QY 61 AAGTGCACGTGTTTCTGTCGCAAAA 88  
|||||  
Db 600 AAGTGCACGTGTTTCTGTCGTAANAAAA 627  
|||||

RESULT 10  
CA431735/c  
LOCUS UI-H-DF0-bet-n-02-0-UI.s1 NCI CGAP DFO Homo sapiens cDNA clone  
DEFINITION UI-H-DF0-bet-n-02-0-UI 3', mRNA sequence.

ACCESSION CA431735  
VERSION CA431735.1 GI:24794461  
KEYWORDS EST.  
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo

REFERENCE 1 (bases 1 to 236)  
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs@mail.nih.gov  
Tissue Procurement: Dr. Jose Mercuende  
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
CDNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Clone distribution information can be obtained from Dr. M. Bento Soares, bento-soares@uiowa.edu  
Seq primer: M13 FORWARD  
POLYA=Yes.

FEATURES  
Location/Qualifiers  
1..236  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="UI-H-DF0-bet-n-02-0-UI"  
/tissue\_type="Subchondral Bone"  
/dev\_stage="Adult"  
/lab\_host="DH10B (Life Technologies)"  
/clone\_lib="NCI-CGAP DFO"  
/note="Organ: Bone; Vector: p7T73-Pac (Pharmacia) with a modified polylinker; Site\_1: EcoR I; Site\_2: Not I; NCI CGAP DFO is a cDNA library containing the following tissue(s): Subchondral Bone. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into p7T73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is GTTAAAGCGTC.  
TAG LIB=UI-H-DF0  
TAG\_SEQ=GTTAAGCGTC"

ORIGIN  
Query Match 90.9%; Score 80; DB 6; Length 236;  
Best Local Similarity 94.3%; Pred. No. 2.5e-16;  
Matches 83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCTCTCACACGCTTTTCTACAAATGCAATCAAT 60  
|||||  
Db 99 CGGAGCCGAGTCCTGTATCAGCCCTTTATCTCTCACACGCTTTTCTACAAATGCAATCAAT 40  
|||||

QY 61 AAGTGCACGTGTTTCTGTCGCAAAA 88  
|||||  
Db 39 AAGTGCACGTGTTTCTGTCGTAANAAAA 12  
|||||

RESULT 11  
AI547782/c  
LOCUS UI-R-C3-sj-g-03-0-UI.s1 UI-R-C3 Rattus norvegicus cDNA clone  
DEFINITION UI-R-C3-sj-g-03-0-UI 3', mRNA sequence.

ACCESSION AI547782  
VERSION AI547782.1 GI:4465270  
KEYWORDS EST.  
SOURCE Rattus norvegicus (Norway rat)  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 243)  
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.  
 TITLE Normalization and subtraction: two approaches to facilitate gene discovery  
 JOURNAL Genome Res. 6 (9), 791-806 (1996)  
 PUBMED 8889548  
 COMMENT Contact: Soares, MB  
 Coordinated Laboratory for Computational Genomics  
 University of Iowa  
 375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA  
 Tel: 319 335 8250  
 Fax: 319 335 9565  
 Email: bento-soares@uiowa.edu  
 The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. cDNA library Preparation: M.B. Soares Lab Clone distribution: clones will be available through Research Genetics (www.resgen.com) This clone is also available through the I.M.A.G.E. Consortium at LLNL (info@image.llnl.gov). IMAGE ID=1768820  
 Seq primer: M13 Forward  
 POLYA=No.

# FEATURES

source Location/Qualifiers  
 1..243  
 /organism="Rattus norvegicus"  
 /mol\_type="mRNA"  
 /strain="Sprague-Dawley"  
 /db\_xref="taxon:10116"  
 /clones="UI-R-C3-aj-g-03-0-UI"  
 /dev\_stage="adult"  
 /lab\_hosts="DH10B (Life Technologies)"  
 /clone\_lib="UI-R-C3"  
 /notes="Vector: p773D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-C3 library is a subtracted library of a series, ultimately derived from a mixture of individually tagged normalized libraries from rat placenta, adult lung, brain, liver, kidney, heart, spleen, ovary, muscle, and 8, 12 and 18-day embryos, after a series of subtractions to reduce the representation of cDNAs from which ESTs had already been generated. The following serially subtracted libraries were generated in this process: UI-R-C3, UI-R-C2p, UI-R-C1, UI-R-C0, UI-R-A1, UI-R-E1. The tag is a string of 3-5 nucleotides present between the Not I site and the oligo-dT track which allows identification of the library of origin of a clone within themixture. The subtracted library (UI-R-C3) was constructed as follows: PCR amplified cDNA inserts from UI-R-C2p clones from which 3' ESTs had been derived was used as a driver in a hybridization with the UI-R-C2p library in the form of single-stranded circles. The remaining single-stranded circles (subtracted library) was purified by hydroxyapatite column chromatography, converted to double-stranded circles and electroporated into DH10B bacteria (Life Technologies) to generate the UI-R-C3 library. This procedure has been previously described (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996)."

# ORIGIN

Query Match 90.9%; Score 80; DB 1; Length 243;  
 Best Local Similarity 94.3%; Pred. No. 2.5e-16;  
 Matches 83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 Qy 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCTCAGCGTTTCTACAATGGCATTCAAT 60  
 Db 96 CGGAGCGGAGTCCTGTATCAGCCCTTTATCTCAGCGTTTCTACAATGGCATTCAAT 37  
 Qy 61 AAAGTCACGCTGTTCTTGGTACAAAA 88  
 Db 36 AAAGTCACGCTGTTCTTGGAAAAA 9  
 RESULT 12

CN310115  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 251)  
 Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J., Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R., Lebkowski,J. and Stanton,L.W.  
 Transcriptionome characterization elucidates signaling networks that control human ES cell growth and differentiation  
 Nat. Biotechnol. 22 (6), 707-716 (2004)  
 15146197  
 Contact: Brandenberger R  
 Regenerative Medicine  
 Geron Corporation  
 230 Constitution Drive, Menlo Park, CA 94025, USA  
 Tel: 650 473 8658  
 Fax: 650 473 7760  
 Email: rbrandenberger@geron.com  
 Insert Length: 251 Std Error: 0.00.  
 Location/Qualifiers  
 1..251  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /tissue\_type="embryonic stem cell, retinoic acid and mitogen-treated hES cell line H7"  
 /clone\_lib="GRN\_PRENEU"  
 /notes="Oligo dT primed, full-length enriched cDNA library from hES cell line H7 (p29) maintained in feeder-free conditions. Embryoid bodies were generated in the presence of all-trans retinoic and mitogens."  
 ORIGIN  
 Query Match 90.9%; Score 80; DB 7; Length 251;  
 Best Local Similarity 94.3%; Pred. No. 2.6e-16;  
 Matches 83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 Qy 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCTCAGCGTTTCTACAATGGCATTCAAT 60  
 Db 162 CGGAGCGGAGTCCTGTATCAGCCCTTTATCTCAGCGTTTCTACAATGGCATTCAAT 221  
 Qy 61 AAAGTCACGCTGTTCTTGGTACAAAA 88  
 Db 222 AAAGTCACGCTGTTCTTGGTAAAAA 249  
 RESULT 13  
 AA542950/c  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 271)  
 NCI-CGAP  
 http://www.ncbi.nlm.nih.gov/ncicgap.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
 Unpublished (1997)

CN310115  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 251)  
 Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J., Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R., Lebkowski,J. and Stanton,L.W.  
 Transcriptionome characterization elucidates signaling networks that control human ES cell growth and differentiation  
 Nat. Biotechnol. 22 (6), 707-716 (2004)  
 15146197  
 Contact: Brandenberger R  
 Regenerative Medicine  
 Geron Corporation  
 230 Constitution Drive, Menlo Park, CA 94025, USA  
 Tel: 650 473 8658  
 Fax: 650 473 7760  
 Email: rbrandenberger@geron.com  
 Insert Length: 251 Std Error: 0.00.  
 Location/Qualifiers  
 1..251  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /tissue\_type="embryonic stem cell, retinoic acid and mitogen-treated hES cell line H7"  
 /clone\_lib="GRN\_PRENEU"  
 /notes="Oligo dT primed, full-length enriched cDNA library from hES cell line H7 (p29) maintained in feeder-free conditions. Embryoid bodies were generated in the presence of all-trans retinoic and mitogens."  
 ORIGIN  
 Query Match 90.9%; Score 80; DB 7; Length 251;  
 Best Local Similarity 94.3%; Pred. No. 2.6e-16;  
 Matches 83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 Qy 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCTCAGCGTTTCTACAATGGCATTCAAT 60  
 Db 162 CGGAGCGGAGTCCTGTATCAGCCCTTTATCTCAGCGTTTCTACAATGGCATTCAAT 221  
 Qy 61 AAAGTCACGCTGTTCTTGGTACAAAA 88  
 Db 222 AAAGTCACGCTGTTCTTGGTAAAAA 249  
 RESULT 13  
 AA542950/c  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 271)  
 NCI-CGAP  
 http://www.ncbi.nlm.nih.gov/ncicgap.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
 Unpublished (1997)

CN310115  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 251)  
 Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J., Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R., Lebkowski,J. and Stanton,L.W.  
 Transcriptionome characterization elucidates signaling networks that control human ES cell growth and differentiation  
 Nat. Biotechnol. 22 (6), 707-716 (2004)  
 15146197  
 Contact: Brandenberger R  
 Regenerative Medicine  
 Geron Corporation  
 230 Constitution Drive, Menlo Park, CA 94025, USA  
 Tel: 650 473 8658  
 Fax: 650 473 7760  
 Email: rbrandenberger@geron.com  
 Insert Length: 251 Std Error: 0.00.  
 Location/Qualifiers  
 1..251  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /tissue\_type="embryonic stem cell, retinoic acid and mitogen-treated hES cell line H7"  
 /clone\_lib="GRN\_PRENEU"  
 /notes="Oligo dT primed, full-length enriched cDNA library from hES cell line H7 (p29) maintained in feeder-free conditions. Embryoid bodies were generated in the presence of all-trans retinoic and mitogens."  
 ORIGIN  
 Query Match 90.9%; Score 80; DB 7; Length 251;  
 Best Local Similarity 94.3%; Pred. No. 2.6e-16;  
 Matches 83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 Qy 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCTCAGCGTTTCTACAATGGCATTCAAT 60  
 Db 162 CGGAGCGGAGTCCTGTATCAGCCCTTTATCTCAGCGTTTCTACAATGGCATTCAAT 221  
 Qy 61 AAAGTCACGCTGTTCTTGGTACAAAA 88  
 Db 222 AAAGTCACGCTGTTCTTGGTAAAAA 249  
 RESULT 13  
 AA542950/c  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 271)  
 NCI-CGAP  
 http://www.ncbi.nlm.nih.gov/ncicgap.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
 Unpublished (1997)



COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: M. Bento Soares, Ph.D.  
cDNA Library Arrayed by: Greg Lennon, Ph.D.  
Cloning Sequencing by: Washington University Genome Sequencing Center  
Cloning Distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: [www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)  
Insert Length: 370 Std Error: 0.00  
Seq primer: 40ml3 fwd ET from Amersham.  
Location/Qualifiers

## FEATURES

source

1. 271  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:984252"  
/sex="male"  
/tissue type="normal prostate"  
/lab host="DH10B"  
/clone lib="NCI CGAP Pr21"  
/note="Organ: prostate; Vector: pRT3D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from normal prostate bulk tissue, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pRT3 vector. Library is not normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo."

## ORIGIN

Query Match 90.9%; Score 80; DB 1; Length 271;  
Best Local Similarity 94.3%; Pred. No. 2.6e-16;  
Matches 83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
  
Qy 1 CGGAGTCAGTCCTGTATCAGCCCTTTATCTCACACGCTTTCTACAAATGGCATTCAAT 60  
|||  
Db 88 CGGAGCCGAGTCCTGTATCAGCCCTTTATCTCACACGCTTTCTACAAATGGCATTCAAT 29  
|||  
  
Qy 61 AAAGTGCACGCTTTCTTGGTACAAAA 88  
|||  
Db 28 AAAGTGCACGCTTTCTTGGTACAAAA 1  
|||

RESULT 14  
CN483092  
LOCUS hw28c11.y2 Human primary human ocular pericytes. Unamplified (hw) EST 26-APR-2004  
DEFINITION Homo sapiens cDNA clone hw28c11 5', mRNA sequence.  
ACCESSION CN483092  
VERSION CN483092.1 GI:46564596  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 280)  
Tsai,J.Y. and Wistow,G.  
Expressed sequence tag analysis of cultured primary human ocular pericytes  
Unpublished (2004)  
Contact: Wistow G  
Section on Molecular Structure and Function  
National Eye Institute  
6/331, NIH, Bethesda, MD 20892-2740, USA  
Tel: 301 402 3452  
Fax: 301 496 0078  
Email: graeme@helix.nih.gov  
Plate: 28 row: c column: 11  
Seq primer: M13RP1 reverse primer (ABI).  
Location/Qualifiers

1. 280  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="hw28c11"  
/cell\_type="pericytes"  
/dev\_stage="Adult"  
/lab\_host="EMDH10B"  
/clone lib="Human primary human ocular pericytes.  
Unamplified (hw)"  
/note="Organ: Eye; Vector: pSport1; RNA was extracted from primary human pericytes in culture. A directionally cloned cDNA library in the pSPORT1 vector (Invitrogen) was constructed at Bioserve Biotechnology (Laurel MD) essentially following the protocols of the SuperScript plasmid System full details of which are contained in the manufacturer's instruction manual (<http://www.lifetech.com/>). First strand synthesis was carried out using a Not I primer-adaptor [5'-pGACTAGTTCTAGATCGAGCGGCCG(T)15-3']. cDNA was cloned in Not I/Sal I sites. EST analysis was performed at the NIH Intramural Sequencing Center (NISC)."

## ORIGIN

Query Match 90.9%; Score 80; DB 7; Length 280;  
Best Local Similarity 94.3%; Pred. No. 2.6e-16;  
Matches 83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
  
Qy 1 CGGAGTCAGTCCTGTATCAGCCCTTTATCTCACACGCTTTCTACAAATGGCATTCAAT 60  
|||  
Db 185 CGGAGCCGAGTCCTGTATCAGCCCTTTATCTCACACGCTTTCTACAAATGGCATTCAAT 244  
|||  
  
Qy 61 AAAGTGCACGCTTTCTTGGTACAAAA 88  
|||  
Db 245 AAAGTGCACGCTTTCTTGGTACAAAA 272  
|||

RESULT 15  
BI060666/c  
LOCUS BI060666 313 bp mRNA linear EST 15-JUN-2001  
DEFINITION PMO-UT0103-300101-002-b09 UT0103 Homo sapiens cDNA, mRNA sequence.  
ACCESSION BI060666  
VERSION BI060666.1 GI:14468193  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 313)  
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W.Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., deOliveira,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.  
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
10737800  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: [asimpson@ludwig.org.br](mailto:asimpson@ludwig.org.br)  
This sequence was derived from the FAPESP/LICR Human Cancer Genome project. This entry can be seen in the following URL (<http://www.ludwig.org.br/scripts/gethtml2.pl?cl=PM0&t2=PM0-UT0103-300101-002-b09&t3=2001-01-30&t4=1>)  
Seq primer: puc 18 forward



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 21:43:03 ; Search time 2856.57 Seconds  
(without alignments)  
6427.437 Million cell updates/sec

Title: US-09-980-046B-6  
Perfect score: 323  
Sequence: 1 cggggccagcctatgcct.....gtgcaagtgtgtgaaaaa 323

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues  
Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_in.\*
- 3: gb\_env.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pt.\*
- 9: gb\_ro.\*
- 10: gb\_sts.\*
- 11: gb\_by.\*
- 12: gb\_un.\*
- 13: gb\_vi.\*
- 14: gb\_btg.\*
- 15: gb\_pl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	321.4	99.5	683	CS113008	Sequence
2	321.4	99.5	683	CS113019	Sequence
3	321.4	99.5	851	BC000897	Homo sapi
4	321.4	99.5	851	BD222216	Human nuc
5	321.4	99.5	851	AX014337	Sequence
6	319.8	99.0	1421	CQ412786	Sequence
7	319.4	98.9	1095	AX281854	Sequence
8	318.4	98.6	877	CQ894717	Sequence
9	317.4	98.3	641	CQ726860	Sequence
10	317.4	98.3	21478	CQ869625	Sequence
11	317.4	98.3	143779	AC136475	Homo sapi
12	317.4	98.3	156819	AP006283	Homo sapi
13	317.4	98.3	175416	AP006286	Homo sapi
14	317.4	98.3	187160	AC138230	Homo sapi
15	315.8	97.8	647	CQ776390	Sequence
16	315.8	97.8	647	CQ833999	Sequence
17	315.8	97.8	647	CQ981463	Sequence
18	315.8	97.8	647	CS030890	Sequence

19	315.8	97.8	647	6	CS033357	Sequence
20	315.8	97.8	647	6	CS036047	Sequence
21	315.8	97.8	647	6	CS039842	Sequence
22	315.8	97.8	647	6	CS042309	Sequence
23	315.8	97.8	647	6	CS044999	Sequence
24	315.8	97.8	647	6	CS108186	Sequence
25	315.8	97.8	647	6	CS108354	Sequence
26	315.8	97.8	647	6	CS119483	Sequence
27	315.8	97.8	647	8	HSRNA927	H. sapiens m
28	315.8	97.8	677	6	CQ768913	Sequence
29	315.8	97.8	842	6	AX397528	Sequence
30	315.8	97.8	853	6	CQ812316	Sequence
31	315.8	97.8	853	6	CQ895390	Sequence
32	315.8	97.8	853	6	CS119626	Sequence
33	315.8	97.8	853	6	AX302539	Sequence
34	315.8	97.8	853	6	AX409461	Sequence
35	315.8	97.8	853	6	HUM927A	Human inter
36	314.4	97.3	675	10	BV178634	sqm10170
37	311.8	96.5	764	6	AR148104	Sequence
38	311.8	96.5	764	6	BD084479	Compositi
39	311.8	96.5	764	6	AR350926	Sequence
40	311.8	96.5	764	6	AR433302	Sequence
41	311.8	96.5	764	6	AR615046	Sequence
42	311.8	96.5	764	6	AR642088	Sequence
43	311.8	96.5	764	6	AX316946	Sequence
44	302.2	93.6	427	6	CQ407796	Sequence
45	296.8	91.9	535	6	CQ923236	Sequence

ALIGNMENTS

RESULT 1 CS113008 683 bp DNA linear PAT 24-JUN-2005  
LOCUS Sequence 26 from Patent WO2005054507.  
DEFINITION CS113008  
ACCESSION CS113008  
VERSION CS113008.1 GI:68224580

KEYWORDS :  
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

REFERENCE 1

AUTHORS Corfe, B. and Chirakkal, H.

TITLE Gene screen

JOURNAL Patent: WO 2005054507-A 26 16-JUN-2005;

University of Sheffield (GB)

FEATURES Location/Qualifiers

source 1..683

/organism="Homo sapiens"

/mol\_type="unassigned DNA"

/db\_xref="taxon:9606"

ORIGIN

Query Match	99.5%	Score 321.4	DB 6	Length 683
Best Local Similarity	99.7%	Pred. No. 3.5e-85		
Matches 322	Conservative	0	Mismatches 1	Indels 0
Gaps	0			
QY	1	CGGGGCCCGAGCCCTATGCCTCCACGCCCAAGTGCCTGAACATCTGGGCCCTGATTCGGG	60	
Db	349	CGGGGCCCGAGCCCTATGCCTCCACGCCCAAGTGCCTGAACATCTGGGCCCTGATTCGGG	408	
QY	61	CATCCTCATGACCAATGGATTACCCCTGTTACTGTGTTATTCGGCTCTGTGACAGTCTACCA	120	
Db	409	CATCCTCATGACCAATGGATTACCCCTGTTACTGTGTTATTCGGCTCTGTGACAGTCTACCA	468	
QY	121	TATTATGTTACAGATAACAGGAAAACGGGGTTACTAGTAGCCGCCCATAGCTCGCAA	180	
Db	469	TATTATGTTACAGATAACAGGAAAACGGGGTTACTAGTAGCCGCCCATAGCTCGCAA	528	
QY	181	CTTTTGCACTCCACTGTGCAATGCTGGGCCCTGCACGCTGGGGCTGTGCCCCCTGCCCT	240	

Db	529	CTTTTGCACTCCACTGTGCAATGCTGGCCCTGCACGCTGGGGCTGTTGGCCCTGCCCCCT	588
Qy	241	TGTCCTCGCCCTAGATACAGCAGTTTATATACCCACACACCTGTCTACAGTGTCAATCAAT	300
Db	589	TGTCCTCGCCCTAGATACAGCAGTTTATATACCCACACACCTGTCTACAGTGTCAATCAAT	648
Qy	301	AAAGTGCACGTGCTTGTGAAAAA	323
Db	649	AAAGTGCACGTGCTTGTGAAAAA	671
RESULT 2			
LOCUS	CS113019	683 bp DNA	linear PAT 24-JUN-2005
DEFINITION	Sequence 37 from Patent WO2005054507.		
ACCESSION	CS113019		
VERSION	CS113019.1	GI:68224591	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1		
AUTHORS	Corfe, B. and Chirakkal, H.		
TITLE	Gene screen		
JOURNAL	Patent: WO 2005054507-A 37 16-JUN-2005;		
	University of Sheffield (GB)		
FEATURES	Location/Qualifiers		
source	1..683		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
ORIGIN			
	Query Match 99.5%; Score 321.4; DB 6; Length 683;		
	Best Local Similarity 99.7%; Pred. No. 3.5e-85;		
	Matches 322; Conservative 0; Mismatches 1; Indels 0; Gaps 0;		
Qy	1	CGGGGCCAGGCTATGCTCCACGGCAAGTGCCTGAACATCTGGGCCCTGATTCGGG	60
Db	349	CGGGGCCAGGCTATGCTCCACGGCAAGTGCCTGAACATCTGGGCCCTGATTCGGG	408
Qy	61	CATCCTCATGACCAATGGATTACCCCTGTTACTGGTATTGGCTCTGTGACGTCTACCA	120
Db	409	CATCCTCATGACCAATGGATTATCCTCTGTTACTGTTATTGGCTCTGTGACGTCTACCA	468
Qy	121	TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCTGCAA	180
Db	469	TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCTGCAA	528
Qy	181	CTTTTGCACTCCACTGTGCAATGCTGGCCCTGACGCTGGGGCTGTTGCCCTGCCCCCT	240
Db	529	CTTTTGCACTCCACTGTGCAATGCTGGCCCTGACGCTGGGGCTGTTGCCCTGCCCCCT	588
Qy	241	TGGTCCTGCCCTAGATACAGCAGTTTATATACCCACACACCTGTCTACAGTGTCAATCAAT	300
Db	589	TGGTCCTGCCCTAGATACAGCAGTTTATATACCCACACACCTGTCTACAGTGTCAATCAAT	648
Qy	301	AAAGTGCACGTGCTTGTGAAAAA	323
Db	649	AAAGTGCACGTGCTTGTGAAAAA	671
RESULT 3			
LOCUS	BC000897	683 bp mRNA	linear PRI 29-JUN-2004
DEFINITION	Homo sapiens interferon induced transmembrane protein 1 (9-27), mRNA (cDNA clone MGC:5195 IMAGE:3464598), complete cds.		
ACCESSION	BC000897		
VERSION	BC000897.1	GI:12654158	
KEYWORDS	MGC.		

SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 683)
AUTHORS	Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, P.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, F., Max, S.I., Wang, J., Haieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Young, A.C., Shevchenko, Y., Sanchez, A., Whiting, M., Madan, A., Touchman, J.W., Green, E.D., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Kravinsky, M.I., Skalska, U., Smailus, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBMED	12477932
REFERENCE	2 (bases 1 to 683)
AUTHORS	Strausberg, R.
TITLE	Direct Submission
JOURNAL	Submitted (15-NOV-2000) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK	NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a>
COMMENT	Contact: MGC help desk Email: <a href="mailto:cgapbs-r@mail.nih.gov">cgapbs-r@mail.nih.gov</a> Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLNL) DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305 Web site: <a href="http://www-ahgc.stanford.edu">http://www-ahgc.stanford.edu</a> Contact: (Dickson, Mark) <a href="mailto:medpaxil.stanford.edu">medpaxil.stanford.edu</a> Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Series: IRAK Plate: 4 Row: b Column: 24 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4504580. Location/Qualifiers 1..683 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="MGC:5195 IMAGE:3464598" /tissue_type="Cervix, carcinoma" /clone_lib="NIH MGC_12" /lab_host="DH10B" /note="Vector: pCMV-SPORT6" 1..683 /gene="IFITM1" /note="synonyms: 9-27, LEU13, IFI17, CD225" /db_xref="GeneID:8519" /db_xref="MIM:604456" 131..508 /gene="IFITM1" /codon_start=1 /product="interferon induced transmembrane protein 1 (9-27)"
source	
gene	
CDS	

AX014337	AX014337	851 bp	DNA	linear	PAT 07-SEP-2000
LOCUS	Sequence 46 from Patent WO954353.				
DEFINITION	AX014337				
ACCESSION	AX014337.1	GI:10040691			
VERSION					
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;				
	Hominidae; Homo.				
REFERENCE	1				
AUTHORS	Schmitt,A., Specht,T., Dahl,E., Hinzmann,B., Rosenthal,A. and Pilarczyk,C.				
TITLE	Human nucleic acid sequences of normal uterus tissue				
JOURNAL	Patent: WO 954353-A 46 28-OCT-1999;				
	SCHMITT ARMIN (DE); SPECHT THOMAS (DE); DAHL EDGAR (DE); HINZMANN BERND (DE); ROSENTHAL ANDRE (DE); METAGEN GES FUER GENOMFORSCHUN (DE); PILARSKY CHRISTIAN (DE)				
FEATURES	Location/Qualifiers				
source	1..851				
	/organism="Homo sapiens"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:9606"				
ORIGIN					
Query Match	99.5%;	Score 321.4;	DB 6;	Length 851;	
Best Local Similarity	99.7%;	Pred. No. 3.6e-85;			
Matches 322;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0	
QY	1	CGGGGGCCGCGCTATCGCTCCACCGCCAGTGCCTGAACATCTGGGCCCTGATTCTCGG	60		
Db	500	CGGGGGCCGAGGCCCTATCGCTCCACCGCCAGTGCCTGAACATCTGGGCCCTGATTCTCGG	559		
QY	61	CATCCTCATGACCAATTGGATTCACCCTGTGTTACTGGTATTCGGCTCTGTGACAGTCTACCA	120		

```
Db 560 CATCTCATGACCATGGATTATCTCTGTTACTGGTATTGGCTCTGTGACAGTCTACCA 619
Qy 121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCCGCCCATAGCCTGCAA 180
Db 620 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCCGCCCATAGCCTGCAA 679
Qy 181 CCTTTGCACTCCACTGTCGAATGCTGGCCCTGCACGCTGGGGCTGTTGCCCTGCCCCCT 240
Db 680 CCTTTGCACTCCACTGTCGAATGCTGGCCCTGCACGCTGGGGCTGTTGCCCTGCCCCCT 739
Qy 241 TGGTCTCTGCCCTAGATACAGCAGTTTATATCCACACACACCTGTCTACAGTGTCAATCAAT 300
Db 740 TGGTCTCTGCCCTAGATACAGCAGTTTATATCCACACACACCTGTCTACAGTGTCAATCAAT 799
Qy 301 AAAGTGCACGTGCTTGGA AAAA 323
Db 800 AAAGTGCACGTGCTTGGA AAAA 822

RESULT 6
LOCUS CO412786 1421 bp DNA linear PAT 23-JAN-2004
DEFINITION Sequence 19857 from Patent WO0170979.
ACCESSION CO412786
VERSION CO412786.1 GI:41320567
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE 1
AUTHORS Lee, J. and Lillie, J.
TITLE Genes, compositions, kits, and method for identification,
assessment, prevention, and therapy of ovarian cancer
JOURNAL Patent: WO 0170979-A 19857 27-SEP-2001;
Millennium Pharmaceuticals, Inc. (US)
FEATURES
source Location/Qualifiers
1 .1421
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 99.0%; Score 319.8; DB 6; Length 1421;
Best Local Similarity 99.4%; Pred. No. 1.2e-84;
Matches 321; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 CGGGGCCAGCCCTATGCCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCGGG 60
Db 420 CGGGGCCAGCCCTATGCCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCGGG 479
Qy 61 CATCTCATGACCATGGATTACCCCTGTACTGGTATTGGGCTCTGTGACAGTCTACCA 120
Db 480 CATCTCATGACCATGGATTACCCCTGTACTGGTATTGGGCTCTGTGACAGTCTACCA 539
Qy 121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCCGCCCATAGCCTGCAA 180
Db 540 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCCGCCCATAGCCTGCAA 599
Qy 181 CCTTTGCACTCCACTGTCGAATGCTGGCCCTGCACGCTGGGGCTGTTGCCCTGCCCCCT 240
Db 600 CCTTTGCACTCCACTGTCGAATGCTGGCCCTGCACGCTGGGGCTGTTGCCCTGCCCCCT 659
Qy 241 TGGTCTCTGCCCTAGATACAGCAGTTTATATCCACACACACCTGTCTACAGTGTCAATCAAT 300
Db 660 TGGTCTCTGCCCTAGATACAGCAGTTTATATCCACACACACCTGTCTACAGTGTCAATCAAT 719
Qy 301 AAAGTGCACGTGCTTGGA AAAA 323
Db 720 AAAGTGCACGTGCTTGGA AACA 742
```

```
RESULT 7
LOCUS AX281854 1095 bp DNA linear PAT 02-NOV-2001
DEFINITION Sequence 263 from Patent WO0177389.
ACCESSION AX281854
VERSION AX281854.1 GI:16609105
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE 1
AUTHORS Shiffman, D., Somogyi, R., Lawn, R., Seilhamer, J. J., Porter, G. J.,
Mikita, T. and Tai, J.
TITLE Genes expressed in foam cell differentiation
JOURNAL Patent: WO 0177389-A 263 18-OCT-2001;
Incyte Genomics, Inc. (US)
FEATURES
source Location/Qualifiers
1 .1095
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/ncfe="Incyte ID No: 988231.7"

ORIGIN
Query Match 98.9%; Score 319.4; DB 6; Length 1095;
Best Local Similarity 99.7%; Pred. No. 1.5e-84;
Matches 320; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 CGGGGCCAGCCCTATGCCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCGGG 60
Db 775 CGGGGCCAGCCCTATGCCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCGGG 834
Qy 61 CATCTCATGACCATGGATTACCCCTGTACTGGTATTGGGCTCTGTGACAGTCTACCA 120
Db 835 CATCTCATGACCATGGATTACCCCTGTACTGGTATTGGGCTCTGTGACAGTCTACCA 894
Qy 121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCCGCCCATAGCCTGCAA 180
Db 895 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCCGCCCATAGCCTGCAA 954
Qy 181 CCTTTGCACTCCACTGTCGAATGCTGGCCCTGCACGCTGGGGCTGTTGCCCTGCCCCCT 240
Db 955 CCTTTGCACTCCACTGTCGAATGCTGGCCCTGCACGCTGGGGCTGTTGCCCTGCCCCCT 1014
Qy 241 TGGTCTCTGCCCTAGATACAGCAGTTTATATCCACACACACCTGTCTACAGTGTCAATCAAT 300
Db 1015 TGGTCTCTGCCCTAGATACAGCAGTTTATATCCACACACACCTGTCTACAGTGTCAATCAAT 1074
Qy 301 AAAGTGCACGTGCTTGGA AAA 321
Db 1075 AAAGTGCACGTGCTTGGA AAA 1095

RESULT 8
LOCUS CQ894717 877 bp DNA linear PAT 05-NOV-2004
DEFINITION Sequence 27 from Patent EP1471075.
ACCESSION CQ894717
VERSION CQ894717.1 GI:55467466
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE 1
AUTHORS Rosenthal, A. D., Pilarsky, C., Dahl, E., Specht, T., Bruemendorf, T.,
Lichtner, R., Staub, E., Roepcke, S. and Li, X. I.
TITLE Human nucleic acid sequences expressed in pancreatic carcinomas
JOURNAL Patent: EP 1471075-A 27 27-OCT-2004;
```

```
FEATURES
  source
    Hinzmamm, Bernd (DE); Rosenthal, Andre (DE); Pilarsky, Christian
    (DE); Dahl, Edgar (DE); Specht, Thomas (DE); Lichtner, Rosemarie
    (DE)
    Location/Qualifiers
      1..877
      /organism="Homo sapiens"
      /mol_type="unassigned DNA"
      /db_xref="taxon:9606"
ORIGIN
  Query Match      98.6%; Score 318.4; DB 6; Length 877;
  Best Local Similarity 99.7%; Pred. No. 2.9e-84;
  Matches 319; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
  Qy 1 CGGGGCCAGGCTATGCTCCACGCGCAAGTGCTGAACATCTGGGCCCTGATTCGGG 60
  Db 558 CGGGGCCAGGCTATGCTCCACGCGCAAGTGCTGAACATCTGGGCCCTGATTCGGG 617
  Qy 61 CATCTCATGACCATTTGGATTACCCCTGTTACTGGTATTTCGGCTCTGTGACAGTCTACCA 120
  Db 618 CATCTCATGACCATTTGGATTATCTCTGTTACTGTTATTCGGCTCTGTGACAGTCTACCA 677
  Qy 121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCCGCCATAGCCTGCAA 180
  Db 678 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCCGCCATAGCCTGCAA 737
  Qy 181 CCTTTGCACTCCACTGTGCAATGTGCGCCCTGCACGCTGGGGCTGTGGCCCTGCCCT 240
  Db 738 CCTTTGCACTCCACTGTGCAATGTGCGCCCTGCACGCTGGGGCTGTGGCCCTGCCCT 797
  Qy 241 TGGTCTGCCCTTAGATACAGAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 300
  Db 798 TGGTCTGCCCTTAGATACAGAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 857
  Qy 301 AAAGTGCAAGTGTGTGAA 320
  Db 858 AAAGTGCAAGTGTGTGAA 877
  RESULT 9
  LOCUS      CQ726860      641 bp      DNA      linear      PAT 03-FEB-2004.
  DEFINITION Sequence 12794 from Patent WO02068579.
  ACCESSION CQ726860
  VERSION CQ726860.1 GI:42291172
  KEYWORDS
  SOURCE      Homo sapiens (human)
  ORGANISM    Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
              Hominidae; Homo.
  REFERENCE  1
  AUTHORS    Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
  TITLE      Kites, such as nucleic acid arrays, comprising a majority of
              humanexons or transcripts, for detecting expression and other uses
              thereof
  JOURNAL    Patent: WO 02068579-A 12794 06-SEP-2002;
              PE Corporation (NY) (US)
  FEATURES
    source
      1..641
      /organism="Homo sapiens"
      /mol_type="unassigned DNA"
      /db_xref="taxon:9606"
ORIGIN
  Query Match      98.3%; Score 317.4; DB 6; Length 641;
  Best Local Similarity 99.7%; Pred. No. 5.5e-84;
  Matches 318; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
  Qy 1 CGGGGCCAGGCTATGCTCCACGCGCAAGTGCTGAACATCTGGGCCCTGATTCGGG 60
  Db 323 CGGGGCCAGGCTATGCTCCACGCGCAAGTGCTGAACATCTGGGCCCTGATTCGGG 382
```

```
Qy 61 CATCTCATGACCATTTGGATTACCCCTGTTACTGGTATTTCGGCTCTGTGACAGTCTACCA 120
Db 383 CATCTCATGACCATTTGGATTATCTCTGTTACTGGTATTTCGGCTCTGTGACAGTCTACCA 442
Qy 121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCCGCCATAGCCTGCAA 180
Db 443 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCCGCCATAGCCTGCAA 502
Qy 181 CCTTTGCACTCCACTGTGCAATGTGCGCCCTGCACGCTGGGGCTGTGGCCCTGCCCT 240
Db 503 CCTTTGCACTCCACTGTGCAATGTGCGCCCTGCACGCTGGGGCTGTGGCCCTGCCCT 562
Qy 241 TGGTCTGCCCTTAGATACAGAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 300
Db 563 TGGTCTGCCCTTAGATACAGAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 622
Qy 301 AAAGTGCAAGTGTGTGAA 319
Db 623 AAAGTGCAAGTGTGTGAA 641
RESULT 10
LOCUS      CQ869625      21478 bp      DNA      linear      PAT 13-SEP-2004.
DEFINITION Sequence 46 from Patent WO2004074320.
ACCESSION CQ869625
VERSION CQ869625.1 GI:51999484
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
              Hominidae; Homo.
REFERENCE  1
AUTHORS    Morris, D.W., Morris, D.W. and Malandro, M.S.
TITLE      Novel therapeutic targets in cancer
JOURNAL    Patent: WO 2004074320-A 46 02-SEP-2004;
              Sagres Discovery, Inc. (US)
FEATURES
  source
    1..21478
    /organism="Homo sapiens"
    /mol_type="unassigned DNA"
    /db_xref="taxon:9606"
ORIGIN
  Query Match      98.3%; Score 317.4; DB 6; Length 21478;
  Best Local Similarity 99.7%; Pred. No. 8.6e-84;
  Matches 318; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
  Qy 1 CGGGGCCAGGCTATGCTCCACGCGCAAGTGCTGAACATCTGGGCCCTGATTCGGG 60
  Db 17039 CGGGGCCAGGCTATGCTCCACGCGCAAGTGCTGAACATCTGGGCCCTGATTCGGG 17098
  Qy 61 CATCTCATGACCATTTGGATTACCCCTGTTACTGGTATTTCGGCTCTGTGACAGTCTACCA 120
  Db 17099 CATCTCATGACCATTTGGATTATCTCTGTTACTGGTCTGTGACAGTCTACCA 17158
  Qy 121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCCGCCATAGCCTGCAA 180
  Db 17159 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCCGCCATAGCCTGCAA 17218
  Qy 181 CCTTTGCACTCCACTGTGCAATGTGCGCCCTGCACGCTGGGGCTGTGGCCCTGCCCT 240
  Db 17219 CCTTTGCACTCCACTGTGCAATGTGCGCCCTGCACGCTGGGGCTGTGGCCCTGCCCT 17278
  Qy 241 TGGTCTGCCCTTAGATACAGAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 300
  Db 17279 TGGTCTGCCCTTAGATACAGAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 17338
  Qy 301 AAAGTGCAAGTGTGTGAA 319
  Db 17339 AAAGTGCAAGTGTGTGAA 17357
```





```

/rpt_family="AT_rich"
4592..4901
/rpt_family="AluX"
4902..4968
/rpt_family="AT_rich"
4979..5091
/rpt_family="L1MB4a"
5154..5251
/rpt_family="L1MB4a"
5620..5683
/rpt_family="L1MB4a"
complement(5684..5954)
/rpt_family="AluJo"
5955..6050
/rpt_family="L1MB4a"
complement(6059..6206)
/rpt_family="MER46C"
6349..6578
/rpt_family="L1MB4a"
6792..7203
/rpt_family="L2"
complement(6884..9052)
/rpt_family="L1MC5"
9225..9287
/rpt_family="MER68-int"
9676..9986
/rpt_family="AluSg"
9994..10032
/rpt_family="AT_rich"
10182..10417
/rpt_family="L1MB3"
complement(10418..10946)
/rpt_family="LTR59"
complement(10947..11251)
/rpt_family="AluX"
complement(11252..11272)
/rpt_family="LTR59"
complement(11290..11595)
/rpt_family="AluSp"
11702..12014
/rpt_family="AluSg"
12047..12137
/rpt_family="L1M4"
complement(12159..12465)
/rpt_family="AluX"
12466..12605
/rpt_family="L2"
12763..12832
/rpt_family="L2"
12898..13715
/rpt_family="L1MB5"
13716..14007
/rpt_family="AluSc"
14018..14320
/rpt_family="AluX"
14321..14603
/rpt_family="L1MB5"

```

Query Match 98.3%; Score 317.4; DB 8; Length 143779;  
Best Local Similarity 99.7%; Pred. No. 1.1e-83;  
Matches 318; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Qy 1 CGGGCCCGGCTATGCTTCCACCGCAAGTGCCTGAACTCTGGGCCCTGATTCGGG 60
Db 99498 CGGGCCCGGCTATGCTTCCACCGCAAGTGCCTGAACTCTGGGCCCTGATTCGGG 99557

Qy 61 CATCTCTATGACCATTTGATTCACCTCTGTTACTGTATTTCGGCTGTGACAGTCTACCA 120
Db 99558 CATCTCTATGACCATTTGATTCACCTCTGTTACTGTATTTCGGCTGTGACAGTCTACCA 99617

Qy 121 TATTATGTTACAGATAAATACAGAAAAACGGGGTTTACTAGTAGCGGCCCATAGCCTGCAA 180
Db 99618 TATTATGTTACAGATAAATACAGAAAAACGGGGTTTACTAGTAGCGGCCCATAGCCTGCAA 99677

```

```

Qy 181 CCTTTGCACTCCACTGTGCAATGCTGCGCCCTGCACGCTGGGGCTGTGCCCTGCCCTC 240
Db 99678 CCTTTGCACTCCACTGTGCAATGCTGCGCCCTGCACGCTGGGGCTGTGCCCTGCCCTC 99737

Qy 241 TGGTCTCGCCCTAGATACAGCAGTTTATATACCCACACACACCTGTCTACAGTGTCAAT 300
Db 99738 TGGTCTCGCCCTAGATACAGCAGTTTATATACCCACACACACCTGTCTACAGTGTCAAT 99797

Qy 301 AAGTGCACTGCTTGGA 319
Db 99798 AAGTGCACTGCTTGGA 99816

```

```

RESULT 12
AP006283/c 156819 bp DNA linear HTG 26-MAR-2003
LOCUS Homo sapiens chromosome 11 clone RP11-326C3 map 1lp, WORKING DRAFT
DEFINITION SEQUENCE, 19 unordered pieces.
ACCESSION AP006283
VERSION AP006283.1 GI:29243343
KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE Homo sapiens genomic DNA of 1lp
JOURNAL Published Only in Database (2003)
REFERENCE 2
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (24-MAR-2003) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
----- Genome Center
Center: RIKEN Genomic Sciences Center(GSC)
Center code: RIKEN
Web site: http://hgp.gsc.riken.go.jp/
Contact: hattori@gsc.riken.go.jp
----- Project Information
Center project name: HumDraft11
Center clone name: RP11-326C3
----- Summary Statistics
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 144706 bases at least Q40
Consensus quality: 150967 bases at least Q30
Consensus quality: 153657 bases at least Q20
Insert size: 155019; sum-of-contigs
Quality coverage: 6.36x in Q20 bases; sum-of-contigs
-----

```

NOTE: This is a 'working draft' sequence. It currently consists of 19 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved

```

1 21001 contig of 21001 bp in length
21102 40714 contig of 19613 bp in length
40815 58430 contig of 17616 bp in length
58531 66754 contig of 8224 bp in length
66855 81134 contig of 14280 bp in length
81235 89763 contig of 8529 bp in length
89864 101889 contig of 12026 bp in length

```

101990 112625 contig of 10636 bp in length  
112726 contig of 7077 bp in length  
125181 contig of 5279 bp in length  
129184 contig of 3903 bp in length  
132282 contig of 7976 bp in length  
137260 contig of 6182 bp in length  
137361 contig of 3249 bp in length  
143643 contig of 2577 bp in length  
146891 contig of 3010 bp in length  
149588 contig of 1357 bp in length  
149669 contig of 1427 bp in length  
152779 contig of 1057 bp in length  
154135 contig of 1057 bp in length  
154236 contig of 1057 bp in length  
155662 contig of 1057 bp in length  
155763 contig of 1057 bp in length.  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 19 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1  
21001: contig of 21001 bp in length  
21101: Gap of 100 bp  
40714: contig of 19613 bp in length  
40814: Gap of 100 bp  
58430: contig of 17616 bp in length  
58530: Gap of 100 bp  
66754: contig of 8224 bp in length  
66854: Gap of 100 bp  
81134: contig of 14280 bp in length  
81234: Gap of 100 bp  
85763: contig of 8529 bp in length  
89863: Gap of 100 bp  
101889: contig of 12026 bp in length  
101890: Gap of 100 bp  
112625: contig of 10636 bp in length  
112725: Gap of 100 bp  
119802: contig of 7077 bp in length  
119902: Gap of 100 bp  
125181: contig of 5279 bp in length  
125281: Gap of 100 bp  
129184: contig of 3903 bp in length  
129284: Gap of 100 bp  
132282: contig of 7976 bp in length  
137260: Gap of 100 bp  
137361: contig of 3249 bp in length  
143643: Gap of 100 bp  
146891: contig of 3010 bp in length  
149588: contig of 1357 bp in length  
149669: Gap of 100 bp  
152779: contig of 6182 bp in length  
154135: Gap of 100 bp  
154236: contig of 1057 bp in length  
155662: contig of 1057 bp in length  
155763: Gap of 100 bp  
155763: contig of 1057 bp in length.

FEATURES  
source

1. .156819  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/map="11p"  
/chromosome="11"  
/clone="RP11-326C3"  
1. .21001  
/note="assembly\_fragment"  
21102. .40714  
/note="assembly\_fragment"  
40815. .58430  
/note="assembly\_fragment"  
58531. .66754  
/note="assembly\_fragment clone\_end:SP6 vector\_side:left"

misc\_feature 66855. .81134  
/note="assembly\_fragment"  
misc\_feature 81235. .89763  
/note="assembly\_fragment"  
misc\_feature 89864. .101889  
/note="assembly\_fragment"  
misc\_feature 101990. .112625  
/note="assembly\_fragment"  
misc\_feature 112726. .119802  
/note="assembly\_fragment"  
misc\_feature 119903. .125181  
/note="assembly\_fragment"  
misc\_feature 125282. .129184  
/note="assembly\_fragment clone\_end:T7 vector\_side:right"  
misc\_feature 129285. .137260  
/note="assembly\_fragment"  
misc\_feature 137361. .143542  
/note="assembly\_fragment"  
misc\_feature 143643. .146891  
/note="assembly\_fragment"  
misc\_feature 146992. .149568  
/note="assembly\_fragment"  
misc\_feature 149669. .152678  
/note="assembly\_fragment"  
misc\_feature 152779. .154135  
/note="assembly\_fragment"  
misc\_feature 154236. .155662  
/note="assembly\_fragment"  
misc\_feature 155763. .156819  
/note="assembly\_fragment"  
  
ORIGIN  
  
Query Match 98.3%; Score 317.4; DB 14; Length 156819;  
Best Local Similarity 99.7%; Pred. No. 1.1e-83;  
Matches 318; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 1 CGGGGGCCAGGCTATGCTCCACCGCCAGTGGTAACTCTGGGCCCTGATTCGGG 60  
Db 77394 CGGGGGCCAGGCTATGCTCCACCGCCAGTGGTAACTCTGGGCCCTGATTCGGG 77335  
  
Qy 61 CATCCTCATGACCATTCGATTCACCCCTGTTACTGGTATTCTGGCTCTGTGACAGTCTACCA 120  
Db 77334 CATCCTCATGACCATTCGATTCACCCCTGTTACTGGTATTCTGGCTCTGTGACAGTCTACCA 77275  
  
Qy 121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTACCGCCCACTAGCTCGCAA 180  
Db 77274 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTACCGCCCACTAGCTCGCAA 77215  
  
Qy 181 CCTTGCACCTCCACTGTGCAATGCTGGCCCTGCACGCTGGGGCTGTTGGCCCTGCCCT 240  
Db 77214 CCTTGCACCTCCACTGTGCAATGCTGGCCCTGCACGCTGGGGCTGTTGGCCCTGCCCT 77155  
  
Qy 241 TGGTCTCTGCCCTAGATACAGCAGTTTATATACCCACACACCTGTCTACAGTGTCTTCAAT 300  
Db 77154 TGGTCTCTGCCCTAGATACAGCAGTTTATATACCCACACACCTGTCTACAGTGTCTTCAAT 77095  
  
Qy 301 AAAGTGCACGTGCTGTGGA 319  
Db 77094 AAAGTGCACGTGCTGTGGA 77076

RESULT 13  
AP006286/c  
LOCUS  
DEFINITION Homo sapiens chromosome 11 clone RP13-25N22 map 11p, WORKING DRAFT  
SEQUENCE, 22 unordered pieces.  
ACCESSION AP006286  
VERSION AP006286.1 GI:29243346  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

```

REFERENCE
1
AUTHORS
Hattori,M., Ishii,K., Toyoda A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE
Homo sapiens genomic DNA of 1lp
JOURNAL
Published Only in Database (2003)
REFERENCE
2
(bases 1 to 175416)
AUTHORS
Hattori,M., Ishii,K., Toyoda A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE
Direct Submission
JOURNAL
Submitted (24-MAR-2003) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Sushiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
----- Genomic Center
Center: RIKEN Genomic Sciences Center (GSC)
Center code: RIKEN
Web site: http://hgp.gsc.riken.go.jp/
Contact: hattori@gsc.riken.go.jp
----- Project Information
Center project name: HumDraft11
Center clone name: RP13-25N22
----- Summary Statistics
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 160074 bases at least Q40
Consensus quality: 177722 bases at least Q30
Consensus quality: 171662 bases at least Q20
Insert size: 173316; sum-of-contigs
Quality coverage: 7.92x in Q20 bases; sum-of-contigs
-----
NOTE: This is a 'working draft' sequence. It currently consists of
22 contigs. The true order of the pieces is not known and their
order in this sequence record is arbitrary. Gaps between the
contigs are represented as runs N, but the exact sizes of the gaps
are unknown. This record will be updated with the finished sequence
as soon as it is available and the accession number will be
preserved.
1
33131 contig of 33131 bp in length
33232 54098 contig of 20867 bp in length
54199 72037 contig of 17839 bp in length
72137 83277 contig of 11140 bp in length
72138 83277 contig of 11140 bp in length
83378 98599 contig of 15222 bp in length
98700 105513 contig of 6814 bp in length
105614 116229 contig of 10616 bp in length
116330 121964 contig of 5635 bp in length
122065 127627 contig of 4261 bp in length
127228 131971 contig of 4244 bp in length
132072 137506 contig of 5435 bp in length
137607 143829 contig of 6223 bp in length
143930 148539 contig of 4610 bp in length
148640 153148 contig of 4509 bp in length
153249 157509 contig of 4261 bp in length
157510 160561 contig of 2952 bp in length
160562 164778 contig of 4117 bp in length
164779 168269 contig of 3391 bp in length
168270 168369 contig of 100 bp
168370 170689 contig of 2320 bp in length
170690 170789 contig of 100 bp
170790 172833 contig of 2044 bp in length
172834 174155 contig of 1222 bp in length
174156 174255 contig of 100 bp
174256 175416 contig of 1161 bp in length.

FEATURES
source
1. .175416
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="1lp"
/clone="RP13-25N22"
1. .33131
/note="assembly_fragment clone_end:SP6 vector_side:left"
33232..54098
/note="assembly_fragment"
54199..72037
/note="assembly_fragment"
72138..83277
/note="assembly_fragment"
83378..98599
/note="assembly_fragment"
98700..105513
/note="assembly_fragment clone_end:T7 vector_side:left"
105614..116229
/note="assembly_fragment"
116330..121964
/note="assembly_fragment"
122065..127627
/note="assembly_fragment"
127728..131971
/note="assembly_fragment"
132072..137506
/note="assembly_fragment"
137607..143829
/note="assembly_fragment"
143930..148539
/note="assembly_fragment"
misc_feature
1. .33131
/note="assembly_fragment clone_end:SP6 vector_side:left"
33232..54098
/note="assembly_fragment"
54199..72037
/note="assembly_fragment"
72138..83277
/note="assembly_fragment"
83378..98599
/note="assembly_fragment"
98700..105513
/note="assembly_fragment clone_end:T7 vector_side:left"
105614..116229
/note="assembly_fragment"
116330..121964
/note="assembly_fragment"
122065..127627
/note="assembly_fragment"
127728..131971
/note="assembly_fragment"
132072..137506
/note="assembly_fragment"
137607..143829
/note="assembly_fragment"
143930..148539
/note="assembly_fragment"

```

```

Hominidae; Homo.
1
Hattori,M., Ishii,K., Toyoda A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE
Homo sapiens genomic DNA of 1lp
JOURNAL
Published Only in Database (2003)
REFERENCE
2
(bases 1 to 175416)
AUTHORS
Hattori,M., Ishii,K., Toyoda A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE
Direct Submission
JOURNAL
Submitted (24-MAR-2003) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Sushiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
----- Genomic Center
Center: RIKEN Genomic Sciences Center (GSC)
Center code: RIKEN
Web site: http://hgp.gsc.riken.go.jp/
Contact: hattori@gsc.riken.go.jp
----- Project Information
Center project name: HumDraft11
Center clone name: RP13-25N22
----- Summary Statistics
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 160074 bases at least Q40
Consensus quality: 177722 bases at least Q30
Consensus quality: 171662 bases at least Q20
Insert size: 173316; sum-of-contigs
Quality coverage: 7.92x in Q20 bases; sum-of-contigs
-----
NOTE: This is a 'working draft' sequence. It currently consists of
22 contigs. The true order of the pieces is not known and their
order in this sequence record is arbitrary. Gaps between the
contigs are represented as runs N, but the exact sizes of the gaps
are unknown. This record will be updated with the finished sequence
as soon as it is available and the accession number will be
preserved.
1
33131 contig of 33131 bp in length
33232 54098 contig of 20867 bp in length
54199 72037 contig of 17839 bp in length
72137 83277 contig of 11140 bp in length
72138 83277 contig of 11140 bp in length
83378 98599 contig of 15222 bp in length
98700 105513 contig of 6814 bp in length
105614 116229 contig of 10616 bp in length
116330 121964 contig of 5635 bp in length
122065 127627 contig of 4261 bp in length
127228 131971 contig of 4244 bp in length
132072 137506 contig of 5435 bp in length
137607 143829 contig of 6223 bp in length
143930 148539 contig of 4610 bp in length
148640 153148 contig of 4509 bp in length
153249 157509 contig of 4261 bp in length
157510 160561 contig of 2952 bp in length
160562 164778 contig of 4117 bp in length
164779 168269 contig of 3391 bp in length
168270 168369 contig of 100 bp in length
168370 170689 contig of 2320 bp in length
170690 170789 contig of 100 bp in length
170790 172833 contig of 2044 bp in length
172834 174155 contig of 1222 bp in length
174156 174255 contig of 100 bp in length
174256 175416 contig of 1161 bp in length.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 22 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
33131: contig of 33131 bp in length
33232 33231: gap of 100 bp
54098: contig of 20867 bp in length

```

```
/note="assembly_fragment"
148640..153148
/note="assembly_fragment"
153249..157509
/note="assembly_fragment"
157610..160561
/note="assembly_fragment"
160662..164778
/note="assembly_fragment"
164879..168269
/note="assembly_fragment"
168370..170689
/note="assembly_fragment"
170790..172833
/note="assembly_fragment"
172934..174155
/note="assembly_fragment"
174256..175416
/note="assembly_fragment"

ORIGIN
Query Match      98.3%; Score 317.4; DB 14; Length 175416;
Best Local Similarity 99.7%; Pred. No. 1.1e-83;
Matches 318; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGGCCCGCCCTATCGCTCCACCGCCAAAGTGGCTGAAACATCTGGGCGCTGATCTCGG 60
Db 75568 CGGGCCCGCCCTATCGCTCCACCGCCAAAGTGGCTGAAACATCTGGGCGCTGATCTCGG 75509

Qy 61 CATCCTCATGACCATTCGATTCACCCCTGTTACTGGTATTCGGCTCTGTGACAGTCTACCA 120
Db 75508 CATCCTCATGACCATTCGATTCATCCTGTTACTGGTATTCGGCTCTGTGACAGTCTACCA 75449

Qy 121 TATTATGTTACAGATAATACAGAAAAACGGGGTTACTAGTAGCCGCCCATAGCCTGCAA 180
Db 75448 TATTATGTTACAGATAATACAGAAAAACGGGGTTACTAGTAGCCGCCCATAGCCTGCAA 75389

Qy 181 CCTTTGCACCTCCTGTCATGTCGCTGACGCTGACGCTGGGCTGTGGCCCTGCCCCCT 240
Db 75388 CCTTTGCACCTCCTGTCATGTCGCTGACGCTGGGCTGTGGCCCTGCCCCCT 75329

Qy 241 TGGTCTCCCTCCTAGATACAGCAGTTTATATACCCACACACCTGTCTACAGTGTCAATCAAT 300
Db 75328 TGGTCTCCCTCCTAGATACAGCAGTTTATATACCCACACACCTGTCTACAGTGTCAATCAAT 75269

Qy 301 AAGTGCACGCTGCTTGGA 319
Db 75268 AAGTGCACGCTGCTTGGA 75250

RESULT 14
AC138230
LOCUS AC138230 187160 bp DNA linear PRI 22-MAY-2003
DEFINITION Homo sapiens chromosome 11, clone RP13-317D12, complete sequence.
ACCESSION AC138230
VERSION AC138230.5 GI:30984765
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 187160)
AUTHORS Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barna,N., Bactien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B.,
Camarata,J., Chang,J., Chazaro,K., Choepel,Y., Collymore,A.,
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., Mactean,C.,
Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
Meldrim,J., Meneus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J.,
Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
Rachupka,A., Ramasaamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C.,
Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
Talamas,J., Testaye,S., Theodore,J., Topham,K., Travers,M.,
Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (20-DEC-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 187160)
REFERENCE 4 (bases 1 to 187160)
AUTHORS Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
Collymore,A., Cook,A., Cooke,P., Corum,B., Dearellano,K.,
Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., Mactean,C.,
Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
Meldrim,J., Meneus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J.,
Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
Rachupka,A., Ramasaamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C.,
Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
Talamas,J., Testaye,S., Theodore,J., Topham,K., Travers,M.,
Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (10-APR-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 187160)
REFERENCE 4 (bases 1 to 187160)
AUTHORS Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
Collymore,A., Cook,A., Cooke,P., Corum,B., Dearellano,K.,
Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., Mactean,C.,
Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
Meldrim,J., Meneus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J.,
Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
Rachupka,A., Ramasaamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C.,
Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
Talamas,J., Testaye,S., Theodore,J., Topham,K., Travers,M.,
Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (22-MAY-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 22, 2003 this sequence version replaced gi:29423936.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
```

Center project name: L29013  
Center clone name: 317\_D\_12

-----  
FEATURES  
Source Location/Qualifiers  
1..187160  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="11"  
/map="11"  
/clone="RP13-317D12"  
135..200  
/clone\_lib="RP13-317D12 Human Female BAC"  
134..1310  
/rpt\_family="L2"  
134..1310  
/rpt\_family="MIR"  
complement(1311..1614)  
/rpt\_family="AluY"  
1615..1638  
/rpt\_family="MIR"  
1849..1938  
/rpt\_family="L1ME"  
complement(4297..4385)  
/rpt\_family="L2"  
4463..4561  
/rpt\_family="FRAM"  
complement(4573..4613)  
/rpt\_family="MAD81"  
4733..5040  
/rpt\_family="AluY"  
complement(5094..5383)  
/rpt\_family="L2"  
complement(5379..5505)  
/rpt\_family="L2"  
complement(5501..5627)  
/rpt\_family="L2"  
complement(5667..6096)  
/rpt\_family="L2"  
6542..6666  
/rpt\_family="MER41E"  
6699..6719  
/rpt\_family="GC-rich"  
6917..6996  
/rpt\_family="C-rich"  
7136..7329  
/rpt\_family="C-rich"  
15461..15483  
/rpt\_family="TGn"  
18178..18413  
/rpt\_family="L2"  
complement(18543..18907)  
/rpt\_family="MER4A"  
complement(18913..19138)  
/rpt\_family="AluSg/x"  
complement(19139..19318)  
/rpt\_family="MER4D"  
complement(19319..19688)  
/rpt\_family="MER4E1"  
19689..19992  
/rpt\_family="AluSp"  
19994..20308  
/rpt\_family="AluSg"  
complement(20309..20630)  
/rpt\_family="MER4E1"  
20632..20763  
/rpt\_family="AluSg/x"  
complement(20785..20882)  
/rpt\_family="MER4E"  
complement(20886..21091)  
/rpt\_family="MER41E"  
21092..21402  
/rpt\_family="AluSx"  
complement(21403..21580)

repeat\_region /rpt\_family="MER41E"  
complement(22261..22538)  
repeat\_region /rpt\_family="MER41E"  
complement(22543..22594)  
repeat\_region /rpt\_family="MER4D"  
22595..22902  
/rpt\_family="AluY"  
repeat\_region complement(22903..23195)  
/rpt\_family="MER4D"  
23197..23314  
/rpt\_family="FLAM C"  
repeat\_region complement(23318..23368)  
/rpt\_family="MER4D"  
23369..23890  
/rpt\_family="MER4A"  
repeat\_region complement(23891..24166)  
/rpt\_family="MER4D"  
repeat\_region complement(24381..24698)  
/rpt\_family="MER63"  
repeat\_region complement(25372..25668)  
/rpt\_family="AluY"  
repeat\_region 26402..26493  
/rpt\_family="(TG)n"  
repeat\_region complement(27325..28044)  
/rpt\_family="LTR30"  
28747..28755  
/notes="<30 qual SNGL region"  
repeat\_region 29024..29047  
/rpt\_family="GC-rich"  
repeat\_region complement(29397..30103)  
/rpt\_family="LTR8"  
repeat\_region 30537..30656  
/rpt\_family="LTR2C"

Query Match 98.3%; Score 317.4; DB 8; Length 187160;  
Best Local Similarity 99.7%; Pred. No. 1.1e-83;  
Matches 318; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGGCCCGAGCCCTATGCTCCACCGCAAGTGCCTGAAACATCTGGGCCCTGATTCGGG 60  
Db 32833 CGGGCCCGAGCCCTATGCTCCACCGCAAGTGCCTGAAACATCTGGGCCCTGATTCGGG 32892  
Qy 61 CATCCTCATGACCATTTGGATTACCCCTGTTACTGGTATTCCGGCTCTGTGACAGTCTACCA 120  
Db 32893 CATCCTCATGACCATTTGGATTACCCCTGTTACTGGTATTCCGGCTCTGTGACAGTCTACCA 32952  
Qy 121 TATTATGTTACAGATAATACAGGAAACCGGGTTACTAGTAGCCGCCCATAGCTGCAA 180  
Db 32953 TATTATGTTACAGATAATACAGGAAACCGGGTTACTAGTAGCCGCCCATAGCTGCAA 33012  
Qy 181 CCTTTGCACTCCACTGTGCAATGCTGGCCCTGCAAGCTGGGGCTGTTGCCCTGCCCT 240  
Db 33013 CCTTTGCACTCCACTGTGCAATGCTGGCCCTGCAAGCTGGGGCTGTTGCCCTGCCCT 33072  
Qy 241 TGGTCTGCCCCCTAGATACAGCAGTTTATATACCAACACACCTGTCTACAGTGTCAATCAAT 300  
Db 33073 TGGTCTGCCCCCTAGATACAGCAGTTTATATACCAACACACCTGTCTACAGTGTCAATCAAT 33132  
Qy 301 AAAGTGACGCTGCTGTGA 319  
Db 33133 AAAGTGACGCTGCTGTGA 33151

RESULT 15  
CQ776390  
LOCUS CQ776390 647 bp DNA linear PAT 11-MAR-2004  
DEFINITION Sequence 76 from Patent EP1394274.  
ACCESSION CQ776390  
VERSION CQ776390.1 GI:45379780  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1  
AUTHORS Ontani,N., Sugita,Y., Yamaya,M., Kubo,H., Nagai,H. and Izuhara,K.  
TITLE Methods of testing for bronchial asthma or chronic obstructive  
JOURNAL pulmonary disease  
Patent: EP 1394274-A 76 03-MAR-2004;  
Genox Research, Inc. (JP)  
FEATURES Location/Qualifiers  
source 1..647  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN

Query Match 97.8%; Score 315.8; DB 6; Length 647;  
Best Local Similarity 99.4%; Pred. No. 1.7e-83;  
Matches 317; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 CGGGGCCCGAGCCCTATGCTCCACCGCAAGTGCCTGAAACATCTGGGCCCTGATTCCTGGG 60  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
329 CGGGGCCCGAGCCCTATGCTCCACCGCAAGTGCCTGAAACATCTGGGCCCTGATTCCTGGG 388  
Qy 61 CATCCTCATGACCAATTGGATTACCCCTGTTACTGGTATTTCGGCTCTGTGACAGTCTACCA 120  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
389 CATCCTCATGACCAATTGGATTACCCCTGTTACTGGTATTTCGGCTCTGTGACAGTCTACCA 448  
Qy 121 TATTATGTTACAGATAATACAGGAAACCGGGTTACTAGTAGCCGCCCATAGCCTGCAA 180  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
449 TATTATGTTACAGATAATACAGGAAACCGGGTTACTAGTAGCCGCCCATAGCCTGCAA 508  
Qy 181 CCTTTGCACCTCCACTGTGCAATGCTGGCCCTGCACGCTGGGGCTGTTGCCCTGCCCCCT 240  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
509 CCTTTGCACCTCCACTGTGCAATGCTGGCCCTGCACGCTGGGGCTGTTGCCCTGCCCCCT 568  
Qy 241 TGGTCTGCCCCCTAGATACAGCAGTTTATACCCACACACACCTGTCTACAGTGTCAATCAAT 300  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
569 TGGTCTGCCCCCTAGATACAGCAGTTTATACCCACACACACCTGTCTACAGTGTCAATCAAT 628  
Qy 301 AAAGTGACGCTGCTTGTA 319  
Db ||||||||||||||||||||  
629 AAAGTGACGCTGCTTGTA 647

Search completed: December 7, 2005, 04:22:04  
Job time : 2859.57 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 11:47:47 ; Search time 507.782 Seconds  
(without alignments)  
4239.411 Million cell updates/sec

Title: US-09-980-046B-6

Perfect score: 323

Sequence: 1 cggggccagcagcctatgcct.....gtgcagctgctgtgaaaaa 323

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N\_Geneseq\_21.\*

1: geneseqn1980s.\*

2: geneseqn1990s.\*

3: geneseqn2000s.\*

4: geneseqn2001as.\*

5: geneseqn2001bs.\*

6: geneseqn2002as.\*

7: geneseqn2002bs.\*

8: geneseqn2003as.\*

9: geneseqn2003bs.\*

10: geneseqn2003cs.\*

11: geneseqn2003ds.\*

12: geneseqn2004as.\*

13: geneseqn2004bs.\*

14: geneseqn2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	323	100.0	323	4 AAC89737	AAC89737 Human gas
2	321.4	99.5	683	14 AEA36105	Aea36105 Human nuc
3	321.4	98.5	583	14 AEA36094	Aea36094 Human nuc
4	321.4	99.5	715	13 ADT50813	Adt50813 Cancer re
5	321.4	99.5	851	2 AA241369	Aa241369 Human nor
6	321.4	99.5	1377	13 ADT50814	Adt50814 Cancer re
7	319.8	99.0	1421	5 ADL45967	Adl45967 Human ova
8	319.4	98.9	1095	6 AAS95008	Aas95008 Human DNA
9	318.4	98.6	877	13 ADX97479	Adx97479 Pancreati
10	318.2	98.5	593	6 ABQ60272	Abq60272 Human col
11	317.4	98.3	1617	10 ADE25685	Ade25685 Human cdn
12	317.4	98.3	21478	13 ABD32562	Abd32562 Human can
13	315.8	97.8	647	6 ABQ60782	Abq60782 Human 9-2
14	315.8	97.8	647	8 ABX10351	Abx10351 DNA encod
15	315.8	97.8	647	10 ADD18965	Add18965 Human dis
16	315.8	97.8	647	12 ADU74824	Adj74824 Marker ge
17	315.8	97.8	647	12 ADL82898	Adl82898 Human PRO
18	315.8	97.8	647	12 ADQ29609	Adq29609 Human col
19	315.8	97.8	647	13 ADR25073	Adr25073 Breast ca

20	315.8	97.8	647	13 ADL14228	Adl14228 Human NF-
21	315.8	97.8	647	13 ADP24680	Adp24680 PRO polyP
22	315.8	97.8	647	13 ADU06094	Adu06094 Novel bro
23	315.8	97.8	647	14 ADY14590	Ady14590 DNA encod
24	315.8	97.8	647	14 ADY19747	Ady19747 DNA encod
25	315.8	97.8	647	14 ADY17057	Ady17057 DNA encod
26	315.8	97.8	647	14 AEA04400	Aea04400 Human CDN
27	315.8	97.8	647	14 AEA23652	Aea23652 Human PRO
28	315.8	97.8	647	14 AEA23820	Aea23820 Human PRO
29	315.8	97.8	677	12 ADI26257	Adi26257 Human Fra
30	315.8	97.8	842	6 ABK46192	Abk46192 cDNA encod
31	315.8	97.8	853	5 ABA83110	Abas83110 IFN-induc
32	315.8	97.8	853	6 ABK83770	Abk83770 Human CDN
33	315.8	97.8	853	6 ABK64497	Abk64497 Human ben
34	315.8	97.8	853	6 ABN95610	Abn95610 Gene #210
35	315.8	97.8	853	10 ADH28927	Adh28927 Human chr
36	315.8	97.8	853	10 ADK60882	Adk60882 Ovarian c
37	315.8	97.8	853	13 ADS64446	Ads64446 Human int
38	315.8	97.8	853	14 ACL54036	ACL54036 Human col
39	315.2	97.6	527	6 ABQ60421	Abq60421 Human molec
40	311.8	96.5	764	2 AAV69033	Aav69033 DNA molec
41	311.8	96.5	764	3 AAC80995	Aac80995 Human bre
42	311.8	96.5	764	6 AAS99841	Aas99841 Breast tu
43	311.8	96.5	764	6 ABK46885	Abk46885 Human bre
44	311.8	96.5	764	8 ADA11362	Ada11362 Human bre
45	311.8	96.5	764	10 ADC15335	Adc15335 Human bre

#### ALIGNMENTS

RESULT 1

AAC89737

ID AAC89737 standard; cDNA; 323 BP.

AC AAC89737;

XX 12-MAR-2001 (first entry)

XX Human gastrointestinal inflammation-related cDNA, SEQ ID NO: 6.

XX Human; cytostatic; immunomodulator; immunostimulant; veterinary;  
anti-inflammatory; neuroprotective; antibacterial; gene therapy;  
gastrointestinal inflammation; immune system disorder; genetic disorder;  
cancer; autoimmune disorder; infection; wound healing; ss.

OS Homo sapiens.

XX WO200073324-A2.

XX 07-DEC-2000.

XX 01-JUN-2000; 2000WO-US015191.

XX 01-JUN-1999; 99US-0137058P.

XX (DIGI-) DIGITAL GENE TECHNOLOGIES INC.

XX Youakim A, Dubose RF, Sims JE, Pribyl TM, Hillbush BS, Haseel KW;

XX WPI; 2001-061508/07.

XX New polynucleotides and polypeptides, useful in gene therapy and in  
diagnosing a pathological condition, e.g. for modulating gene expression  
in gastrointestinal inflammation, or for treating cancers or genetic  
disorders.

XX Claim 1; Page 85; 108pp; English.

XX The present sequence is one of a number of isolated human polynucleotides  
which are useful in gene therapy, and for diagnosing a pathological  
condition or a susceptibility to it. In particular, the polynucleotides  
are useful for modulating gene expression in gastrointestinal

CC inflammation. The polynucleotides are useful for chromosome  
CC identification, controlling gene expression through triple helix  
CC formation or antisense DNA or RNA, or identifying individuals from minute  
CC biological samples using DNA-based identification techniques. The  
CC polynucleotides can also be used as an alternative to restriction  
CC fragment length polymorphism (RFLP), by determining the actual base-by-  
CC base DNA sequences of selected portions of an individual's genome. The  
CC polynucleotides may also be used as molecular weight markers on Southern  
CC gels, as diagnostic probes for the presence of a specific mRNA, as a  
CC probe to subtract-out known sequences in the process of discovering novel  
CC polynucleotides, or as an antigen to elicit an immune response. The  
CC polypeptides are useful in diagnostic procedures to detect a disorder.  
CC The polynucleotides and polypeptides are useful for preventing, treating  
CC or ameliorating immune system disorders, genetic disorders, cancers, some  
CC autoimmune disorders, or infections. The polynucleotides and polypeptides  
CC are also useful for differentiating, proliferating or attracting cells,  
CC leading to the regeneration of tissues, especially in wounds or burns.  
CC The polypeptides and polynucleotides may also be used as a food additive  
CC or preservative

XX  
SQ Sequence 323 BP; 71 A; 99 C; 72 G; 81 T; 0 U; 0 Other;

Query Match 100.0%; Score 323; DB 4; Length 323;  
Best Local Similarity 100.0%; Pred. No. 1.1e-91;  
Matches 323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGGGCCAGGCTATGCTCCACCGCAAGTGCTGAAACATCTGGGCCCTGATTCGGG 60  
|||||  
Db 1 CGGGGCCAGGCTATGCTCCACCGCAAGTGCTGAAACATCTGGGCCCTGATTCGGG 60  
|||||  
Qy 61 CATCCTCATGACCATTCGATTACCCCTGTTACTGGTATTCGGCTCTGTGACAGTCTACCA 120  
|||||  
Db 61 CATCCTCATGACCATTCGATTACCCCTGTTACTGGTATTCGGCTCTGTGACAGTCTACCA 120  
|||||  
Qy 121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCCTGCAA 180  
|||||  
Db 121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCCTGCAA 180  
|||||  
Qy 181 CTTTGTCACTCCACTGTGCAATGTGCGCCCTGCAAGCTGGGGCTGTGGCCCTGCCCCCT 240  
|||||  
Db 181 CTTTGTCACTCCACTGTGCAATGTGCGCCCTGCAAGCTGGGGCTGTGGCCCTGCCCCCT 240  
|||||  
Qy 241 TGGTCTGCCCTAGATACAGCAGTTTATACCCACACACCTGCTACAGTGTCAATCAAT 300  
|||||  
Db 241 TGGTCTGCCCTAGATACAGCAGTTTATACCCACACACCTGCTACAGTGTCAATCAAT 300  
|||||  
Qy 301 AAAGTGACAGTGTGTTGAAAAA 323  
|||||  
Db 301 AAAGTGACAGTGTGTTGAAAAA 323  
|||||

## RESULT 2

AEA36105  
ID AEA36105 standard; DNA; 683 BP.

XX AEA36105;

XX 25-AUG-2005 (first entry)

XX Human nucleic acid sequence #37.

XX Screening; gene expression; colorectal tumor; colitis; Crohns disease;  
XX irritable bowel syndrome; gastrointestinal disease; cytostatic;  
XX gastrointestinal-gen.; antiinflammatory; ds.

XX Homo sapiens.

XX WO2005054507-A2.

XX 16-JUN-2005.

XX 03-DEC-2004; 2004WO-GB005078.

PR 04-DEC-2003; 2003GB-00028048.

XX (UYSH-) UNIV SHEFFIELD.

PA Corfe B, Chirakkal H;

XX WPI; 2005-435407/44.

XX Screening for nucleic acid molecules exhibiting altered expression in  
PT cells grown in the presence of butyrate, and detection of the nucleic  
PT acid molecules or the encoded polypeptides in diagnosing colorectal  
PT cancer.

XX Disclosure; Page 68; 266pp; English.

XX The invention relates to a method of screening for nucleic acid molecules  
CC that show altered expression in a first cell sample comprising comparing  
CC the gene expression profile of the sample with that of a second reference  
CC sample, where the first sample has been grown in the presence of butyrate  
CC or a related carbon source from which butyrate is directly or indirectly  
CC derived, but the reference sample has not. The invention also relates to  
CC a method of detecting at least one nucleic acid molecule associated with  
CC the initiation and/or progression of colorectal cancer in an animal,  
CC comprising providing a biological sample comprising at least one cell to  
CC be tested, contacting the sample with a ligand (preferably a hybridizing  
CC nucleic acid molecule) which binds to at least one nucleic acid and  
CC detecting the presence of at least one molecule in the sample, a method  
CC of detecting at least one polypeptide associated with the initiation  
CC and/or progression of colorectal cancer in an animal comprising providing  
CC a biological sample comprising at least one cell to be tested, contacting  
CC the sample with at least one ligand that specifically binds at least one  
CC polypeptide encoded by a nucleic acid or a variant polypeptide comprising  
CC an amino acid sequence which varies by the addition, deletion or  
CC substitution of at least one amino acid residue and detecting the  
CC presence of the polypeptide in the sample, a method of screening for  
CC agents that modulate the activity of at least one polypeptide encoded by  
CC a gene associated with the initiation and/or progression of colorectal  
CC cancer comprising forming a preparation comprising at least one  
CC polypeptide encoded by a nucleic acid or a variant polypeptide comprising  
CC at least one addition, deletion or substitution and at least one agent to  
CC be tested and determining the activity of the agent with respect to  
CC activity of the polypeptide, and an antibody or its effective binding  
CC portion identified by the method, for use as a pharmaceutical. The  
CC methods are useful for screening for nucleic acid molecules that show  
CC altered expression in a cell sample, and for detecting a nucleic acid and  
CC a polypeptide respectively, that are associated with the initiation  
CC and/or progression of colorectal cancer and are useful for detecting or  
CC monitoring colorectal cancer, especially adenocarcinoma. The methods are  
CC also useful for screening for agents that modulate the activity of at  
CC least one polypeptide encoded by a gene associated with the initiation  
CC and/or progression of cancer, where agents identified by the method are  
CC useful for treating colorectal cancer. The methods could also be used to  
CC detect or monitor other conditions such as colitis, Crohn's disease or  
CC irritable bowel syndrome, as a screening tool for fiber consumption, as  
CC an assay for colon microflora functionality or for early detection of pre  
CC -cancerous growth. This sequence represents a human nucleic acid  
CC identified by the screening method of the invention.

XX  
SQ Sequence 683 BP; 167 A; 214 C; 155 G; 147 T; 0 U; 0 Other;

Query Match 99.5%; Score 321.4; DB 14; Length 683;  
Best Local Similarity 99.7%; Pred. No. 4.9e-91;  
Matches 322; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGGGCCAGGCTATGCTCCACCGCAAGTGCTGAAACATCTGGGCCCTGATTCGGG 60  
|||||  
Db 349 CGGGGCCAGGCTATGCTCCACCGCAAGTGCTGAAACATCTGGGCCCTGATTCGGG 408  
|||||

Qy 61 CATCCTCATGACCATTCGATTACCCCTGTTACTGGTATTCGGCTCTGTGACAGTCTACCA 120  
|||||  
Db 409 CATCCTCATGACCATTCGATTACCCCTGTTACTGGTATTCGGCTCTGTGACAGTCTACCA 468  
|||||  
Qy 121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCCTGCAA 180  
|||||



Db 469 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCTAGCTGCAA 528  
Qy 181 CCTTGGCACTCCACTGTCATGCTGGCCCTGACGCTGGGGCTGTGGCCCTGCCCT 240  
Db 529 CCTTGGCACTCCACTGTCATGCTGGCCCTGACGCTGGGGCTGTGGCCCTGCCCT 588  
Qy 241 TGGTCTGCTCCCTAGATACAGCAGTTTATACCCACACACACCTGTCTACAGTGTCAATCAAT 300  
Db 589 TGGTCTGCTCCCTAGATACAGCAGTTTATACCCACACACACCTGTCTACAGTGTCAATCAAT 648  
Qy 301 AAAGTGCACTGCTGTGAAAAA 323  
Db 649 AAAGTGCACTGCTGTGAAAAA 671

## RESULT 3

AEA36094  
ID AEA36094 standard; DNA; 683 BP.  
AC AEA36094;  
XX  
DT 25-AUG-2005 (first entry)  
XX  
XX Human nucleic acid sequence #26.  
DE  
XX Screening; gene expression; colorectal tumor; colitis; Crohns disease;  
KW Irritable bowel syndrome; gastrointestinal disease; cytostatic;  
KW gastrointestinal-gen.; antiinflammatory; ds.  
XX  
OS Homo sapiens.  
XX  
XX WO2005054507-A2.  
XX  
XX 16-JUN-2005.  
XX  
XX 03-DEC-2004; 2004WO-GB005078.  
XX  
XX 04-DEC-2003; 2003GB-00028048.  
XX  
XX (UYSH-) UNIV SHEFFIELD.  
XX  
XX Corfe B, Chirakkal H;  
XX  
XX WPI; 2005-435407/44.  
DR

Screening for nucleic acid molecules exhibiting altered expression in cells grown in the presence of butyrate, and detection of the nucleic acid molecules or the encoded polypeptides in diagnosing colorectal cancer.

Disclosure; Page 55; 266pp; English.

XX The invention relates to a method of screening for nucleic acid molecules  
XX that show altered expression in a first cell sample comprising comparing  
XX the gene expression profile of the sample with that of a second reference  
XX sample, where the first sample has been grown in the presence of butyrate  
XX or a related carbon source from which butyrate is directly or indirectly  
XX derived, but the reference sample has not. The invention also relates to  
XX a method of detecting at least one nucleic acid molecule associated with  
XX the initiation and/or progression of colorectal cancer in an animal,  
XX comprising providing a biological sample comprising at least one cell to  
XX be tested, contacting the sample with a ligand (preferably a hybridizing  
XX nucleic acid molecule) which binds to at least one nucleic acid and  
XX detecting the presence of at least one molecule in the sample, a method  
XX of detecting at least one polypeptide associated with the initiation  
XX and/or progression of colorectal cancer in an animal comprising providing  
XX a biological sample comprising at least one cell to be tested, contacting  
XX the sample with at least one ligand that specifically binds at least one  
XX polypeptide encoded by a nucleic acid or a variant polypeptide comprising  
XX an amino acid sequence which varies by the addition, deletion or  
XX substitution of at least one amino acid residue and detecting the  
XX presence of the polypeptide in the sample, a method of screening for

CC agents that modulate the activity of at least one polypeptide encoded by  
CC a gene associated with the initiation and/or progression of colorectal  
CC cancer comprising forming a preparation comprising at least one  
CC polypeptide encoded by a nucleic acid or a variant polypeptide comprising  
CC at least one addition, deletion or substitution and at least one agent to  
CC be tested and determining the activity of the agent with respect to  
CC activity of the polypeptide, and an antibody or its effective binding  
CC portion identified by the method, for use as a pharmaceutical. The  
CC methods are useful for screening for nucleic acid molecules that show  
CC altered expression in a cell sample, and for detecting a nucleic acid and  
CC a polypeptide respectively, that are associated with the initiation  
CC and/or progression of colorectal cancer and are useful for detecting or  
CC monitoring colorectal cancer, especially adenocarcinoma. The methods are  
CC also useful for screening for agents that modulate the activity of at  
CC least one polypeptide encoded by a gene associated with the initiation  
CC and/or progression of cancer, where agents identified by the method are  
CC useful for treating colorectal cancer. The methods could also be used to  
CC detect or monitor other conditions such as colitis, Crohn's disease or  
CC irritable bowel syndrome, as a screening tool for fiber consumption, as  
CC an assay for colon microflora functionality or for early detection of pre  
CC -cancerous growth. This sequence represents a human nucleic acid  
CC identified by the screening method of the invention.

XX SQ Sequence 683 BP; 167 A; 214 C; 155 G; 147 T; 0 U; 0 Other;

Query Match 99.5%; Score 321.4; DB 14; Length 683;  
Best Local Similarity 99.7%; Pred. No. 4.9e-91;  
Matches 322; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGGGCCAGGCTATGCTCCACCGCAAGTGCTGAAACATCTGGGCTGTCTGGG 60  
Db 349 CGGGGCCAGGCTATGCTCCACCGCAAGTGCTGAAACATCTGGGCTGTCTGGG 408  
Qy 61 CATCTCATGACCAATTGGATTACCCCTGTTACTGGTATTGGCTCTGTGACAGTCTACCA 120  
Db 409 CATCTCATGACCAATTGGATTACCCCTGTTACTGGTATTGGCTCTGTGACAGTCTACCA 468  
Qy 121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCTAGCTGCAA 180  
Db 469 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCTAGCTGCAA 528  
Qy 181 CCTTTGCACCTCCACTGTGCAATGCTGGCCCTGCAACGCTGGGGCTGTGGCCCTGCCCT 240  
Db 529 CCTTTGCACCTCCACTGTGCAATGCTGGCCCTGCAACGCTGGGGCTGTGGCCCTGCCCT 588  
Qy 241 TGGTCTGCTCCCTAGATACAGCAGTTTATACCCACACACACCTGTCTACAGTGTCAATCAAT 300  
Db 589 TGGTCTGCTCCCTAGATACAGCAGTTTATACCCACACACACCTGTCTACAGTGTCAATCAAT 648  
Qy 301 AAAGTGCACTGCTGTGAAAAA 323  
Db 649 AAAGTGCACTGCTGTGAAAAA 671

## RESULT 4

ADT50813  
ID ADT50813 standard; DNA; 715 BP.

XX AC ADT50813;

XX DT 13-JAN-2005 (first entry)

XX Cancer related nucleic acid sequence #117.

DE ds; gene; cytostatic; gene therapy; vaccine; diagnosis; breast; colon;  
KW lung; ovarian; prostate; cancer.

XX OS Homo sapiens.

XX XX WO2004092338-A2.

XX PD 28-OCT-2004.

XX XX

```
PF 12-APR-2004; 2004WO-US011104.
XX
XX
XX 11-APR-2003; 2003US-0462399P.
PR 01-JUL-2003; 2003US-0484333P.
XX
XX (DIAD-) DIADEXUS INC.
XX
XX Macina RA, Turner LR, Sun Y, Tam A;
XX WPI; 2004-766851/75.
XX
XX New cancer specific nucleic acid (CasNA) molecules, useful for
PT diagnosing, monitoring the presence of, or treating a patient with
PT breast, colon, lung, ovarian, or prostate cancer.
XX
XX Claim 1; SEQ ID NO 117; 891pp; English.
PS
XX The invention relates to an isolated nucleic acid molecule (I)
CC selectively hybridizing to, or comprising at least 95% sequence identity
CC to, any of the 362 nucleotide sequences fully defined in the
CC specification. The nucleic acid molecules and polypeptides are useful for
CC diagnosing, monitoring the presence of, or treating a patient with
CC breast, colon, lung, ovarian, or prostate cancer. This sequence
CC corresponds to a nucleic acid of the invention.
XX
XX Sequence 715 BP; 160 A; 220 C; 176 G; 159 T; 0 U; 0 Other;
SQ
Query Match 99.5%; Score 321.4; DB 13; Length 715;
Best Local Similarity 99.7%; Pred. No. 5e-91;
Matches 322; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CGGGGGCCAGGCTATGCTCCACCGCCAAAGTGCCTGAACATCTGGGCCCTGATTCGGG 60
DB 393 CGGGGGCCAGGCTATGCTCCACCGCCAAAGTGCCTGAACATCTGGGCCCTGATTCGGG 452
QY 61 CATCCTCATGACCATTTGGATTACCCCTGTTACTGGTATTTCGGCTCTGTGACAGTCTACCA 120
DB 453 CATCCTCATGACCATTTGGATTACCCCTGTTACTGGTATTTCGGCTCTGTGACAGTCTACCA 512
QY 121 TATTATGTTACAGATAATACAGGAAAACGGGGTTACTAGTAGCCGCCCATAGCCTGCAA 180
DB 513 TATTATGTTACAGATAATACAGGAAAACGGGGTTACTAGTAGCCGCCCATAGCCTGCAA 572
QY 181 CCTTTGCACTCCACTGTGCAATGTGGCCCTGCACGCTGGGGCTGTTGCCCTGCCCT 240
DB 573 CCTTTGCACTCCACTGTGCAATGTGGCCCTGCACGCTGGGGCTGTTGCCCTGCCCT 632
QY 241 TGGTCTCTGCCCTTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 300
DB 633 TGGTCTCTGCCCTTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 692
QY 301 AAAGTGCACGTGCTTGTGAAAAA 323
DB 693 AAAGTGCACGTGCTTGTGAAAAA 715
RESULT 5
ID AA241369 standard; cDNA; 851 BP.
XX
XX AA241369;
XX
XX 19-JAN-2000 (first entry)
XX Human normal uterus tissue derived cDNA 45.
XX Human; uterus; cancer; treatment; anticancer; cytostatic; gene therapy;
XX EST; expressed sequence tag; ss.
XX Homo sapiens.
XX DE19817946-A1.
XX
```

```
PD 21-OCT-1999.
XX
XX 17-APR-1998; 98DE-01017946.
XX
XX 17-APR-1998; 98DE-01017946.
XX (META-) METAGEN GES GENOMFORSCHUNG MBH.
XX
XX Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;
XX WPI; 1999-591956/51.
XX
XX New nucleic acid sequences expressed in normal uterine tissues, and
PT derived polypeptides, for treatment of uterine cancer and identification
PT of therapeutic agents.
XX
XX Claim 3; Page 108; 154pp; German.
PS
XX This invention describes novel cDNA sequences (A) highly expressed in
CC normal uterine tissue which can have anticancer and cytostatic activity
CC and can be used for gene therapy. (A) are used (i) for recombinant
CC expression of polypeptides (B) and (ii) to isolate complete genes. (B)
CC are used (i) to identify agents suitable for treatment of uterine cancer;
CC (ii) directly for treating this form of cancer (including expression from
CC gene therapy vectors) and (iii) for generation of specific antibodies.
CC (A) are identified by assembling ESTs (expressed sequence tags) from a
CC particular tissue type before comparison of expression patterns. This
CC allows a significantly longer fragment of the gene to be revealed, so
CC should reduce the number of failures associated with the fact that ESTs
CC from different libraries may represent different parts of the same
CC unknown gene, distorting the estimated frequency of occurrence in a
CC particular tissue. AA241325-241385 represent the human uterine tissue
CC derived cDNA fragments of the invention which encode the protein
XX fragments represented in AA259838-Y59892
XX
XX Sequence 851 BP; 225 A; 248 C; 211 G; 167 T; 0 U; 0 Other;
SQ
Query Match 99.5%; Score 321.4; DB 2; Length 851;
Best Local Similarity 99.7%; Pred. No. 5.4e-91;
Matches 322; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CGGGGGCCAGGCTATGCTCCACCGCCAAAGTGCCTGAACATCTGGGCCCTGATTCGGG 60
DB 500 CGGGGGCCAGGCTATGCTCCACCGCCAAAGTGCCTGAACATCTGGGCCCTGATTCGGG 559
QY 61 CATCCTCATGACCATTTGGATTACCCCTGTTACTGGTATTTCGGCTCTGTGACAGTCTACCA 120
DB 560 CATCCTCATGACCATTTGGATTACCCCTGTTACTGGTATTTCGGCTCTGTGACAGTCTACCA 619
QY 121 TATTATGTTACAGATAATACAGGAAAACGGGGTTACTAGTAGCCGCCCATAGCCTGCAA 180
DB 620 TATTATGTTACAGATAATACAGGAAAACGGGGTTACTAGTAGCCGCCCATAGCCTGCAA 679
QY 181 CCTTTGCACTCCACTGTGCAATGTGGCCCTGCACGCTGGGGCTGTTGCCCTGCCCT 240
DB 680 CCTTTGCACTCCACTGTGCAATGTGGCCCTGCACGCTGGGGCTGTTGCCCTGCCCT 739
QY 241 TGGTCTCTGCCCTTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 300
DB 740 TGGTCTCTGCCCTTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 799
QY 301 AAAGTGCACGTGCTTGTGAAAAA 323
DB 800 AAAGTGCACGTGCTTGTGAAAAA 822
RESULT 6
ID ADT50814 standard; DNA; 1377 BP.
XX
XX ADT50814;
XX
XX 13-JAN-2005 (first entry)
XX
```

XX DE Cancer related nucleic acid sequence #118.  
XX DE ds; gene; cytostatic; gene therapy; vaccine; diagnosis; breast; colon;  
XX KW lung; ovarian; prostate; cancer.  
XX KW Homo sapiens.  
XX OS WO2004092338-A2.  
XX PN 28-OCT-2004.  
XX PD 12-APR-2004; 2004WO-US011104.  
XX PF 11-APR-2003; 2003US-0462399P.  
XX PR 01-JUL-2003; 2003US-0484333P.  
XX XX (DIAD-) DIADEXUS INC.  
XX XX Macina RA, Turner LR, Sun Y, Tam A;  
XX XX WPI; 2004-766851/75.  
XX XX New cancer specific nucleic acid (CaSNA) molecules, useful for  
PT diagnosing, monitoring the presence of, or treating a patient with  
PT breast, colon, lung, ovarian, or prostate cancer.  
XX XX Claim 1; SEQ ID NO 118; 891pp; English.  
XX XX The invention relates to an isolated nucleic acid molecule (I)  
CC selectively hybridizing to, or comprising at least 95% sequence identity  
CC to, any of the 362 nucleotide sequences fully defined in the  
CC specification. The nucleic acid molecules and polypeptides are useful for  
CC diagnosing, monitoring the presence of, or treating a patient with  
CC breast, colon, lung, ovarian, or prostate cancer. This sequence  
CC corresponds to a nucleic acid of the invention.  
XX XX Sequence 1377 BP; 346 A; 403 C; 356 G; 263 T; 0 U; 9 Other;  
SQ  
Query Match 99.5%; Score 321.4; DB 13; Length 1377;  
Best Local Similarity 99.7%; Pred. No. 6.8e-91;  
Matches 322; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Oy 1 CGGGCCCGAGGCTATGCTCTCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCTGGG 60  
Db 845 CGGGCCCGAGGCTATGCTCTCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCTGGG 904  
Oy 61 CATCTCTATGACCATTTGGATTACCCCTGTTACTGTTATTCGGCTCTGTGACAGTCTACCA 120  
Db 905 CATCTCTATGACCATTTGGATTACCCCTGTTACTGTTATTCGGCTCTGTGACAGTCTACCA 964  
Oy 121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCTGCAA 180  
Db 965 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCTGCAA 1024  
Oy 181 CTTTGTGACTCCACGTGCAATGCTGGCCCTGTCAGCTGGGCTGTTCGCCCTGCCCT 240  
Db 1025 CTTTGTGACTCCACGTGCAATGCTGGCCCTGTCAGCTGGGCTGTTCGCCCTGCCCT 1084  
Oy 241 TGGTCTCTGCCCTAGATACAGCAGTATTATACCACACACCTGTCTACAGTGTCAATCAAT 300  
Db 1085 TGGTCTCTGCCCTAGATACAGCAGTATTATACCACACACCTGTCTACAGTGTCAATCAAT 1144  
Oy 301 AAGTGCACGGCTTGTGAAAAA 323  
Db 1145 AAGTGCACGGCTTGTGAAAAA 1167  
RESULT 7  
ADL45967  
ID ADL45967 standard; DNA; 1421 BP.  
XX  
AC ADL45967;

XX 20-MAY-2004 (first entry)  
XX Human ovarian cancer DNA marker #19857.  
XX DE Human; ovarian cancer; ds; tumour; cytostatic; DNA marker.  
XX KW Homo sapiens.  
XX OS WO200170979-A2.  
XX PN 27-SEP-2001.  
XX PD 21-MAR-2001; 2001WO-US009126.  
XX PF 21-MAR-2000; 2000US-0191031P.  
XX PR 25-MAY-2000; 2000US-0207124P.  
XX PR 15-JUN-2000; 2000US-0211940P.  
XX PR 07-JUL-2000; 2000US-0216820P.  
XX PR 25-JUL-2000; 2000US-0220661P.  
XX PR 21-DEC-2000; 2000US-0257672P.  
XX XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
XX PA Lee J, Lillie J;  
XX PT WPI; 2001-611502/70.  
XX DR Novel isolated nucleic acid molecules (markers) overexpressed in ovarian  
XX cancer cells as compared to their normal non-cancerous ovarian cells are  
PT used to characterize stage, grade, histological type of ovarian cancer.  
XX Disclosure; SEQ ID NO 19857; 106pp; English.  
XX The invention relates to nucleic acid markers which are overexpressed in  
XX ovarian cancer cells as compared to their expression in normal (i.e. non-  
CC cancerous) ovarian cells. The invention also relates to polypeptides  
CC encoded by the markers, antibodies that selectively bind to the  
CC polypeptides, a method of inhibiting ovarian cancer in a patient at risk  
CC of developing ovarian cancer involving inhibiting expression of a gene  
CC corresponding to a marker of the invention and a method of treating a  
CC patient afflicted with ovarian cancer comprising providing to cells of  
CC the patient an antisense oligonucleotide complementary to a marker of the  
CC invention. The markers are useful for assessing if a patient is afflicted  
CC with ovarian cancer, which involves comparing the level of expression of  
CC a marker in a patient sample and a normal level of expression of the  
CC marker in a control non-ovarian cancer sample. A difference between the  
CC expression levels indicates ovarian cancer. The level of expression of a  
CC marker corresponds to a secreted protein or to a transcribed  
CC polynucleotide or its portion. The level of expression of the marker is  
CC assessed by detecting the presence in the sample, a protein or protein  
CC fragment corresponding to the marker. The presence of protein or protein  
CC fragment is detected using an antibody that specifically binds with the  
CC protein or protein fragment. Alternatively, the level of expression of  
CC the marker is assessed by detecting the presence of a transcribed  
CC polynucleotide which anneals with the marker or anneals with a portion of  
CC the polynucleotide comprising the marker, under stringent conditions. The  
CC marker is also used for monitoring the progression of ovarian cancer in a  
CC patient which involves detecting expression of the marker in a patient  
CC sample at a first point in time, repeating the method at a subsequent  
CC time and comparing the level of expression. The method is carried out  
CC using an ovarian tissue sample. A composition comprising a marker,  
CC polypeptide or antibody of the invention is used to treat ovarian cancer.  
CC This sequence represents a human ovarian cancer DNA marker of the  
CC invention. Note: The sequence data for this patent did not form part of  
CC the printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX SQ Sequence 1421 BP; 364 A; 391 C; 328 G; 338 T; 0 U; 0 Other;  
Query Match 99.0%; Score 319.8; DB 5; Length 1421;  
Best Local Similarity 99.4%; Pred. No. 2.2e-90;  
Matches 321; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGGGGCCAGCCCTATGCTCCACGGCNAAGTGCCTGNAACATCTGGGCCCTGATTCCTGGG 60  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
420 CGGGGCCAGGCCCTATGCTCCACGGCAAGTGCTGAACATCTGGGCCCTGATTCCTGGG 479  
  
QY 61 CATCTCATGACCAATGGATTTCACCCCTGTTACTGGTATTCGGCTCTGTGACAGTCTACCA 120  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
480 CATCTCATGACCAATGGATTTCATCTGTTACTGTTATTCGGCTCTGTGACAGTCTACCA 539  
  
QY 121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCCTGCAA 180  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
540 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCCTGCAA 599  
  
QY 181 CCTTTGCACCTCCACTGTGCAATGCTGGCCCTGCAAGCTGGGGCTGTTGCCCTGCCCCCT 240  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
600 CCTTTGCACCTCCACTGTGCAATGCTGGCCCTGCAAGCTGGGGCTGTTGCCCTGCCCCCT 659  
  
QY 241 TGGTCTCGCCCTAGATACAGCAGTTTATATACCCACACACCTGTCTACAGTGTCAATCAAT 300  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
660 TGGTCTCGCCCTAGATACAGCAGTTTATATACCCACACACCTGTCTACAGTGTCAATCAAT 719  
  
QY 301 AAAGTGACGCTGCTTGTAAGAAA 323  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
720 AAAGTGACGCTGCTTGTAAGAAA 742

## RESULT 8

AAS95008  
ID AAS95008 standard; DNA; 1095 BP.

AC AAS95008;

DT 14-FEB-2002 (first entry)

XX Human DNA sequence #263 expressed during foam cell differentiation.

XX Human; foam cell differentiation; atherosclerosis; cerebral stroke;  
XX cardiovascular disorder; coronary artery disease; gene therapy; ds.

XX Homo sapiens.

XX WO200177389-A2.

XX 18-OCT-2001.

XX 04-APR-2001; 2001WO-US011128.

XX 05-APR-2000; 2000US-0195106P.

XX (INCY-) INCYTE GENOMICS INC.

XX Shiffman D, Somogyi R, Lawn R, Seilhamer JJ, Porter GJ, Mikita T;  
XX Tai J;

XX WPI; 2002-010925/01.

XX Composition useful for diagnosis of conditions, disorders or diseases  
XX associated with atherosclerosis, comprises several polynucleotides that  
XX are differentially expressed in foam cell development.

XX Claim 1; Page 303; 315pp; English.

XX The present invention relates to the isolation of human polynucleotide  
XX sequences that are differentially expressed during foam cell  
XX differentiation. The polynucleotide sequences of the invention or a  
XX composition comprising these polynucleotides are useful as a high  
XX throughput method for detecting altered expression of one or more  
XX polynucleotides in a sample. The polynucleotides can be used in the  
XX diagnosis of disorders associated with foam cell development such as  
XX atherosclerosis, cerebral stroke, and cardiovascular disorders such as  
XX coronary artery disease. The polynucleotide sequences can also be used as  
XX PCR primers and probes. The polynucleotides of the invention are also  
XX useful in gene therapy. AAS94746-AAS95021 represent the human

CC polynucleotide sequences of the invention which are differentially  
CC expressed during foam cell differentiation

XX SQ Sequence 1095 BP; 267 A; 326 C; 268 G; 234 T; 0 U; 0 Other;

Query Match 98.9%; Score 319.4; DB 6; Length 1095;  
Best Local Similarity 99.7%; Pred. No. 2.6e-90;  
Matches 320; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGGGCCAGCCCTATGCTCCACCGCCAAAGTGCCTGAAACATCTGGGCCCTGATTCCTGGG 60

Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
775 CGGGGCCAGCCCTATGCTCCACCGCCAAAGTGCCTGAAACATCTGGGCCCTGATTCCTGGG 834

QY 61 CATCTCATGACCAATGGATTTCACCCCTGTTACTGGTATTCGGCTCTGTGACAGTCTACCA 120

Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
835 CATCTCATGACCAATGGATTTCATCTGTTACTGGTATTCGGCTCTGTGACAGTCTACCA 894

QY 121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCCTGCAA 180

Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
895 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCCTGCAA 954

QY 181 CCTTTGCACCTCCACTGTGCAATGCTGGCCCTGCAAGCTGGGGCTGTTGCCCTGCCCCCT 240

Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
955 CCTTTGCACCTCCACTGTGCAATGCTGGCCCTGCAAGCTGGGGCTGTTGCCCTGCCCCCT 1014

QY 241 TGGTCTCGCCCTAGATACAGCAGTTTATATACCCACACACCTGTCTACAGTGTCAATCAAT 300

Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1015 TGGTCTCGCCCTAGATACAGCAGTTTATATACCCACACACCTGTCTACAGTGTCAATCAAT 1074

QY 301 AAAGTGACGCTGCTTGTAAGAAA 321

Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1075 AAAGTGACGCTGCTTGTAAGAAA 1095

## RESULT 9

ADX97479

ID ADX97479 standard; DNA; 877 BP.

XX ADX97479;

XX 21-APR-2005 (first entry)

XX Pancreatic cancer associated human protein encoding DNA, SEQ ID 27.

XX pancreas tumor; cytostatic; gene; ds.

XX Homo sapiens.

XX EPI471075-A2.

XX 27-OCT-2004.

XX 31-MAR-2004; 2004EP-00090124.

XX 31-MAR-2003; 2003DE-01015834.

XX (HINZ/) HINZMANN B.

XX (ROSE/) ROSENTHAL A.

XX (PILA/) PILARSKY C.

XX (DAHL/) DAHL E.

XX (SPEC/) SPECHT T.

XX (LICH/) LICHTNER R.

XX Rosenthal A, Pilarisky C, Dahl E, Specht T, Bruemendorf T;

XX Lichtner R, Staub E, Roepcke S, Li X;

XX WPI; 2004-768082/76.

XX P-PSDB; ADX97550.

XX New nucleic acid differentially expressed in pancreatic tumor tissue, for  
XX use as diagnostic agents and in screening for therapeutic agents.

XX Claim 1; SEQ ID NO 27; 28pp; German.

XX The invention relates to a novel human nucleic acid sequence of the  
CC pancreas and its encoded protein. The invention further comprises:  
CC proteins and peptides, preferably isolated, that contain a sequence  
CC encoded by the novel nucleic acid; and methods for diagnosis and  
CC treatment of pancreatic cancer, using a substance that inhibits or binds  
CC to the protein or its DNA, including: an antisense oligonucleotide, short  
CC interfering RNA or ribozyme directed against the pancreatic protein, an  
CC organic molecule, particularly having a molecular weight below 5000,  
CC especially 300, that binds to the pancreatic DNA, an aptamer or  
CC (monoclonal) antibody, preferably human or humanized, that binds to the  
CC pancreatic DNA, or an anti-idiotypic antibody raised against the  
CC monoclonal antibody, any of which may be derivatized with a reporter  
CC group, cytotoxic compound, immunostimulant and/or radioisotope. The novel  
CC human pancreatic proteins and their encoding DNA have cytostatic  
CC activity. The novel sequences are useful for inhibiting transcription  
CC and/or expression of genes and proteins associated with pancreatic  
CC cancer. This polynucleotide sequence represents the DNA encoding one of  
CC the novel human pancreatic proteins of the invention. Note: This sequence  
CC is not shown in the specification, it has been electronically downloaded  
CC from a DVD-ROM provided with this specification by the European Patent  
CC Office.

XX Sequence 877 BP; 213 A; 264 C; 219 G; 181 T; 0 U; 0 Other;

Query Match 98.6%; Score 318.4; DB 13; Length 877;  
Best Local Similarity 99.7%; Pred. No. 4.9e-90;  
Matches 319; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGGGCCAGGCTATGCTCCACGGCCAAAGTGCCTGAACATCTGGGCCCTGATTCGGG 60  
Db 558 CGGGGCCAGGCTATGCTCCACGGCCAAAGTGCCTGAACATCTGGGCCCTGATTCGGG 617  
Qy 61 CATCTCATGACCATGTGATTCACCTGTACTGTGTTATTCGGCTCTGTGACAGTCTACCA 120  
Db 618 CATCTCATGACCATGTGATTCACCTGTACTGTGTTATTCGGCTCTGTGACAGTCTACCA 677  
Qy 121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCTGCAG 180  
Db 678 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCTGCAG 737  
Qy 181 CTTTGGCACTCCACTGTGCAATGTGCGCCCTGCACGCTGGGGCTGTGGCCCTGCCCCCT 240  
Db 738 CTTTGGCACTCCACTGTGCAATGTGCGCCCTGCACGCTGGGGCTGTGGCCCTGCCCCCT 797  
Qy 241 TGGTCTCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 300  
Db 798 TGGTCTCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 857  
Qy 301 AAAGTGCACGTGCTTGCGAA 320  
Db 858 AAAGTGCACGTGCTTGCGAA 877

RESULT 10

ABQ60272/c

ID ABQ60272 standard; cDNA; 593 BP.

XX ABQ60272;

AC ABQ60272;

XX 02-AUG-2002 (first entry)

DT Human colon cancer related nucleotide sequence SEQ ID NO:3967.

DE Human; colon cancer; cancer; tissue profiling; forensic; mapping;

XX Human; colon cancer; cancer; tissue profiling; forensic; mapping;

KW genetic analysis; diagnostic; antisense therapy; gene; ss.

XX Homo sapiens.

OS Homo sapiens.

XX WO200229086-A2.

PN 11-APR-2002.

XX 11-APR-2002.

XX 11-APR-2002.

XX 11-APR-2002.

XX 11-APR-2002.

PF 02-OCT-2001; 2001WO-US030732.

XX 02-OCT-2001; 2000US-0237271P.

PR (FARB ) BAYER CORP.

XX Burgess C, Astle JH, Carroll B, Catino TJ, Dwivedi P, Molino GA;

PI Thiagalingam A, Lewis ME;

PI MPI; 2002-426115/45.

XX New isolated nucleic acid that is differentially expressed in cancer

XX tissues useful for determining the presence of colon cancer in a cell or

XX tissue type, and in antisense therapy.

XX Claim 1; Fig 1; 796pp; English.

XX ABQ56306 to ABQ60787 represent isolated nucleic acids (I) differentially

XX expressed in cancer tissues. ABQ56306 to ABQ59004 represent proteins

XX encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be

XX used in antisense therapy. An antibody immunoreactive with a polypeptide

XX encoded by (I) is useful for detecting cancer in a patient sample, and

XX for detecting the presence or absence of a polynucleotide encoded by a

XX nucleic acid which hybridizes to (I) in a cell. A probe/primer derived

XX from (I) can be used for determining the presence of a nucleic acid which

XX hybridizes to (I), and for determining the phenotype of cells in a sample

XX of cells from a patient. (I) is useful for determining the presence of

XX colon cancer in a cell or tissue type, for determining the presence or

XX state of other type of cancer, in antisense therapy, to generate

XX macroarrays on a solid surface, to identify a chromosome on which the

XX corresponding gene resides, and in tissue profiling, forensics, genetic

XX analysis, mapping and diagnostic applications. (I) can be used to raise

XX antibodies, and to screen for peptide analogues and antagonists

XX Sequence 593 BP; 143 A; 122 C; 164 G; 143 T; 0 U; 21 Other;

Qy Query Match 98.5%; Score 318.2; DB 6; Length 593;

Db Best Local Similarity 99.1%; Pred. No. 4.8e-90;

Qy Matches 320; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CGGGGCCAGGCTATGCTCCACGGCCAAAGTGCCTGAACATCTGGGCCCTGATTCGGG 60

Db 342 CGGGGCCAGGCTATGCTCCACGGCCAAAGTGCCTGAACATCTGGGCCCTGATTCGGG 283

Qy 61 CATCTCATGACCATGTGATTCACCTGTACTGTGTTATTCGGCTCTGTGACAGTCTACCA 120

Db 282 CATCTCATGACCATGTGATTCACCTGTACTGTGTTATTCGGCTCTGTGACAGTCTACCA 223

Qy 121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCTGCAG 180

Db 222 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCTGCAG 163

Qy 181 CTTTGGCACTCCACTGTGCAATGTGCGCCCTGCACGCTGGGGCTGTGGCCCTGCCCCCT 240

Db 162 CTTTGGCACTCCACTGTGCAATGTGCGCCCTGCACGCTGGGGCTGTGGCCCTGCCCCCT 103

Qy 241 TGGTCTCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 300

Db 102 TGGTCTCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 43

Qy 301 AAAGTGCACGTGCTTGCGAAAAA 323

Db 42 AAAGTGCACGTGCTTGCGAAAAA 20

XX RESULT 11

XX ADE25685

ID ADE25685 standard; cDNA; 1617 BP.

XX ADE25685

AC ADE25685;

XX 29-JAN-2004 (first entry)

DT 29-JAN-2004 (first entry)

XX 29-JAN-2004 (first entry)

DE Human cDNA differentially expressed in foam cells #89.  
XX Human; ss; differential expression; foam cell; LPS; lipopolysaccharide;  
XX cardiovascular disease; atherosclerosis.  
OS Homo sapiens.  
XX US2003194721-A1.  
XX 16-OCT-2003.  
XX 18-SEP-2002; 2002US-00247671.  
XX 19-SEP-2001; 2001US-0323784P.  
XX (INCY-) INCYTE GENOMICS INC.  
XX Mikita T, Shiffman D, Porter JG, Kaeer MR;  
XX WPI; 2003-875398/81.  
XX Combination containing several polynucleotide that are differentially  
XX expressed in foam cells and complements of the polynucleotides, useful  
XX for diagnosing cardiovascular disease or atherosclerosis.  
XX Claim 1; SEQ ID NO 89; 37pp; English.  
XX The invention relates to a combination comprising several polynucleotides  
XX having any one of 127 sequences (S1) such as the sequence of human  
XX calmodulin gene, human mRNA for KIAA0930 protein, leukotriene A4  
XX hydrolase, human CGI-142 protein mRNA, human K+ channel beat 2 subunit  
XX mRNA, etc., and their complements. The cDNAs are differentially expressed  
XX in LPS (lipopolysaccharide)-treated foam cells. Also included are  
XX obtaining an extended or full length gene from a library of nucleic acid  
XX sequences, an expression vector containing the nucleic acids, a host cell  
XX containing the vector, a purified polypeptide appearing as ADE25750 and  
XX ADE25751, producing a protein by culturing the host cell, and a  
XX composition comprising a purified antibody that specifically binds to the  
XX proteins. The foam cell-expressed nucleic acids are useful for a high  
XX throughput detection of differential expression of one or more  
XX polynucleotides in a sample. The sample is from a subject with  
XX atherosclerosis and comparison with a standard defines early, mid or late  
XX stages of the disorder. The foam cell-expressed nucleic acids are useful  
XX for high throughput screening of a library of molecules or compounds to  
XX identify a ligand which binds a polynucleotide. The library is chosen  
XX from DNA molecules, peptides, proteins and RNA molecules. The protein is  
XX useful for a high throughput screening of library of molecules or  
XX compounds to identify at least one ligand which specifically binds a  
XX protein, for purifying a ligand from a sample for making an antibody. The  
XX foam cell-expressed nucleic acids are useful for diagnosing  
XX cardiovascular disorder. The foam cell-expressed nucleic acids are useful  
XX as elements on a microarray which can be used for detecting related  
XX polynucleotide in a sample, diagnosing cardiovascular disease,  
XX atherosclerosis. The present sequence represents a cDNA whose expression  
XX is upregulated in LPS treated foam cells.  
SQ Sequence 1617 BP; 348 A; 480 C; 445 G; 344 T; 0 U; 0 Other;  
Query Match 98.3%; Score 317.4; DB 10; Length 1617;  
Best Local Similarity 99.7%; Pred. No. 1.3e-89;  
Matches 318; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CGGGGCCAGGCTATGCTTCACCGCAAGTGCTGAAACATCTGGCCCTGATTCCTGGG 60  
DB 773 CGGGGCCAGGCTATGCTTCACCGCAAGTGCTGAAACATCTGGCCCTGATTCCTGGG 832  
QY 61 CATCCTCATGACCAATGGATTACCCCTGTACTGGTATTGGCTCTGTGACAGTCTACCA 120  
DB 833 CATCCTCATGACCAATGGATTATCTGTGTTACTGGTATTGGCTCTGTGACAGTCTACCA 892  
QY 121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCTGCAA 180  
DB 893 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCTGCAA 952

QY 181 CCTTTGCACTCCACTGTGCAATGCTGGCCCTGTCACGCTGGGGCTGTTGCCCTGCCCCCT 240  
DB 953 CCTTTGCACTCCACTGTGCAATGCTGGCCCTGTCACGCTGGGGCTGTTGCCCTGCCCCCT 1012  
QY 241 TGGTCTCTGCCCCCTAGATACAGCAGTATTATATACCCACACACCTGTCTACAGTGTCAAT 300  
DB 1013 TGGTCTCTGCCCCCTAGATACAGCAGTATTATATACCCACACACCTGTCTACAGTGTCAAT 1072  
QY 301 AAAGTGCACGTGCTTGTGA 319  
DB 1073 AAAGTGCACGTGCTTGTGA 1091  
RESULT 12  
ABD32562  
ID ABD32562 standard; DNA; 21478 BP.  
XX  
AC ABD32562;  
XX  
DT 18-NOV-2004 (first entry)  
XX  
DE Human cancer-associated genomic DNA HD14-034.  
XX  
XX Human; ds; cancer-associated protein; gene; cytostatic; cancer;  
XX leukaemia; lymphoma; CAP.  
XX  
OS Homo sapiens.  
XX  
XX WO2004074320-A2.  
XX  
PD 02-SEP-2004.  
XX  
PF 17-FEB-2004; 2004WO-US004730.  
XX  
XX 14-FEB-2003; 2003US-00367094.  
PR 14-MAR-2003; 2003US-00388838.  
PR 15-APR-2003; 2003US-00417375.  
PR 13-JUN-2003; 2003US-00461862.  
PR 15-SEP-2003; 2003US-00663431.  
PR 15-DEC-2003; 2003US-00737318.  
XX  
XX (SAGR-) SAGRES DISCOVERY INC.  
XX  
XX Morris DW, Morris DW, Malandro MS;  
PI  
XX WPI; 2004-652914/63.  
DR  
XX New isolated cancer-associated polynucleotides and polypeptides useful  
PT for diagnosing, preventing or treating cancers, especially lymphoma and  
PT leukemia, or in screening for agents that modulate cancer.  
XX  
XX claim 16; seqid 46; 310pp; English.  
PS  
XX The invention relates to an isolated nucleic acid comprising at least 10  
CC contiguous nucleotides of any of the 233 polynucleotide sequences given  
CC in the specification, or its complement. The nucleic acids encode cancer-  
CC associated proteins. Also included are an expression vector comprising  
CC the isolated nucleic acid cited above, a host cell comprising the above  
CC recombinant nucleic acid or expression vector, a microarray for detecting  
CC a cancer-associated (CA) nucleic acid comprising at least one probe  
CC comprising at least 10 contiguous nucleotides of any of the above-  
CC mentioned nucleotide sequences, an isolated polypeptide (encoded within  
CC an open reading frame of a CA sequence selected from any of the 95  
CC polynucleotide sequences as mentioned in the specification, or its  
CC complement), an isolated antibody, (or its antigen binding fragment) that  
CC binds to the above polypeptide, a hybridoma that produces the above  
CC monoclonal antibody, a pharmaceutical composition comprising the above  
CC antibody and a pharmaceutical excipient, a kit for detecting cancer  
CC cells (comprising the antibody cited above, methods for diagnosing cancer  
CC or for detecting the presence or absence of cancer cells in an  
CC individual, a method for inhibiting growth of cancer cells in an  
CC individual, a method for delivering a therapeutic agent to cancer cells

CC in an individual, an electronic library comprising the above  
CC polynucleotide or polypeptide (or their fragments), methods of screening  
CC for anticancer activity or for a bioactive agent capable of modulating  
CC the activity of a CA protein (CAP), methods for detecting cancer  
CC associated with expression of a polypeptide in a test cell sample, a  
CC method for treating cancers and a method for inhibiting the expression of  
CC CA gene in a cell. The composition and methods are useful for detecting,  
CC diagnosing, preventing and treating cancers, especially lymphoma and  
CC leukaemia. These may also be used in screening for agents that modulate  
CC cancer. The present sequence is a human CAP genomic sequence. Note: The  
CC sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 21478 BP; 3963 A; 4821 C; 5071 G; 3826 T; 0 U; 3797 Other;  
  
Query Match 98.3%; Score 317.4; DB 13; Length 21478;  
Best Local Similarity 99.7%; Pred. No. 4.4e-89;  
Matches 318; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 CGGGGCCAGGCTGCTCCACGCCCAAGTGCCTGAACATCTGGGCCCTGATTCGGG 60  
DB 17039 CGGGGCCAGGCTGCTCCACGCCCAAGTGCCTGAACATCTGGGCCCTGATTCGGG 17098  
QY 61 CATCTCATGACCATTTGATTCACCTCTGTACTGTGTTACTCGGCTCTGTGACAGTCTACCA 120  
DB 17099 CATCTCATGACCATTTGATTCATCTCTGTACTGTGTTACTCGGCTCTGTGACAGTCTACCA 17158  
QY 121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCTGCAC 180  
DB 17159 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCTGCAC 17218  
QY 181 CCTTTGACTCCACTGTGCAATGCTGGCCCTGCACGCTGGGGCTGTGGCCCTGCCCT 240  
DB 17219 CCTTTGACTCCACTGTGCAATGCTGGCCCTGCACGCTGGGGCTGTGGCCCTGCCCT 17278  
QY 241 TGGTCTCTGCCCTAGATACAGCAGTTTATACCCACACACACCTGTCTACAGTGTCTTCAAT 300  
DB 17279 TGGTCTCTGCCCTAGATACAGCAGTTTATACCCACACACACCTGTCTACAGTGTCTTCAAT 17338  
QY 301 AAAGTGCACGCTGTGTGA 319  
DB 17339 AAAGTGCACGCTGTGTGA 17357  
  
RESULT 13  
ABQ60782  
ID ABQ60782 standard; cDNA; 647 BP.  
AC ABQ60782;  
XX  
DT 02-AUG-2002 (first entry)  
XX Human 9-27 cDNA sequence SEQ ID NO:4484.  
XX Human; colon cancer; cancer; tissue profiling; forensic; mapping;  
KW genetic analysis; diagnostic; antisense therapy; gene; ss.  
XX Homo sapiens.  
XX WO200229086-A2.  
XX  
XX 11-APR-2002.  
XX  
XX 02-OCT-2001; 2001WO-US030732.  
XX PF  
XX 02-OCT-2000; 2000US-0237271P.  
XX PR  
XX (FARB ) BAYER CORP.  
XX PA  
XX Burgess C, Astle JH, Carroll E, Catino TV, Dwivedi P, Molino GA;  
XX PI Thiagalingam A, Lewis ME;  
XX

DR WPI; 2002-426115/45.  
DR P-PSDB; ABB78999.  
XX  
PT New isolated nucleic acid that is differentially expressed in cancer  
PT tissues useful for determining the presence of colon cancer in a cell or  
XX tissue type, and in antisense therapy.  
PS Claim 1; Fig 2; 796pp; English.  
XX  
CC ABQ56306 to ABQ60787 represent isolated nucleic acids (I) differentially  
CC expressed in cancer tissues. ABB78993 to ABB79004 represent proteins  
CC encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be  
CC used in antisense therapy. An antibody immunoreactive with a polypeptide  
CC encoded by (I) is useful for detecting cancer in a patient sample, and  
CC for detecting the presence or absence of a polynucleotide encoded by a  
CC nucleic acid which hybridises to (I) in a cell. A probe/primer derived  
CC from (I) can be used for determining the presence of a nucleic acid which  
CC hybridises to (I), and for determining the phenotype of cells in a sample  
CC of cells from a patient. (I) is useful for determining the presence of  
CC colon cancer in a cell or tissue type, for determining the presence or  
CC state of other type of cancer, in antisense therapy, to generate  
CC macroarrays on a solid surface, to identify a chromosome on which the  
CC corresponding gene resides, and in tissue profiling, forensics, genetic  
CC analysis, mapping and diagnostic applications. (I) can be used to raise  
CC antibodies, and to screen for peptide analogues and antagonists  
XX  
SQ Sequence 647 BP; 143 A; 211 C; 145 G; 148 T; 0 U; 0 Other;  
  
Query Match 97.8%; Score 315.8; DB 6; Length 647;  
Best Local Similarity 99.4%; Pred. No. 2.8e-89;  
Matches 317; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 1 CGGGGCCAGGCTGCTCCACGCCCAAGTGCCTGAACATCTGGGCCCTGATTCGGG 60  
DB 329 CGGGGCCAGGCTGCTCCACGCCCAAGTGCCTGAACATCTGGGCCCTGATTCGGG 388  
QY 61 CATCTCATGACCATTTGATTCACCTCTGTACTGTGTTACTCGGCTCTGTGACAGTCTACCA 120  
DB 389 CATCTCATGACCATTTGATTCATCTCTGTACTGTGTTACTCGGCTCTGTGACAGTCTACCA 448  
QY 121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCTGCAC 180  
DB 449 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCTGCAC 508  
QY 181 CCTTTGACTCCACTGTGCAATGCTGGCCCTGCACGCTGGGGCTGTGGCCCTGCCCT 240  
DB 509 CCTTTGACTCCACTGTGCAATGCTGGCCCTGCACGCTGGGGCTGTGGCCCTGCCCT 568  
QY 241 TGGTCTCTGCCCTAGATACAGCAGTTTATACCCACACACACCTGTCTACAGTGTCTTCAAT 300  
DB 569 TGGTCTCTGCCCTAGATACAGCAGTTTATACCCACACACACCTGTCTACAGTGTCTTCAAT 628  
QY 301 AAAGTGCACGCTGTGTGA 319  
DB 629 AAAGTGCACGCTGTGTGA 647  
  
RESULT 14  
ABX10351  
ID ABX10351 standard; DNA; 647 BP.  
XX  
XX ABX10351;  
AC  
XX 28-JAN-2003 (first entry)  
DT  
DE DNA encoding protein differentially regulated in prostate cancer #20.  
XX  
KW Prostate cancer; gene expression; differential regulation;  
KW molecular marker; drug target; cancer detection; cancer diagnosis;  
KW cancer staging; cancer grading; cancer assessing; cancer monitoring;  
XX gene; ds.  
XX Homo sapiens.  
OS



XX WO200281638-A2.  
PN 17-OCT-2002.  
XX 08-APR-2002; 2002WO-US010824.  
XX 06-APR-2001; 2001US-0281731P.  
PR 06-APR-2001; 2001US-0281732P.  
XX (ORIG-) ORIGENE TECHNOLOGIES INC.  
PA Sun Z, Jay G;  
XX WPI; 2003-058520/05.  
XX P-PSDB; ABU07449.  
DR Novel genes which are differentially regulated in prostate cancer, useful  
XX for diagnosing prostate cancer in prostate tissue sample and assessing  
PT therapeutic or preventive intervention in prostate cancer patients.  
XX Claim 1; Page 95; 416pp; English.  
XX The invention describes genes (I) which are differentially regulated in  
CC prostate cancer. (I) is useful for diagnosing a prostate cancer in a  
CC sample comprising prostate tissue, which involves determining the number  
CC of target genes which are differentially-regulated in the sample, where  
CC the number is indicative of the probability that the sample comprises  
CC prostate cancer. (I) is useful for assessing a therapeutic or preventive  
CC intervention in a subject having a prostate cancer, which involves  
CC determining the expression levels in a sample comprising prostate tissue  
CC of target genes which are differentially-regulated in prostate cancer.  
CC Preferably, the expression levels of at least 10 genes are determined.  
CC (I) is also useful for identifying agents that modulate a biological  
CC activity of a polypeptide differentially-regulated in prostate cancer  
CC cells, which involves contacting a polypeptide differentially-regulated  
CC in prostate cancer cells with a test agent under conditions effective for  
CC the test agent to modulate a biological activity of the polypeptide, and  
CC determining whether the test agent modulates the biological activity. (I)  
CC is useful as molecular markers, as drug targets, and for detecting,  
CC diagnosing, staging, grading, assessing, monitoring, prognosticating,  
CC preventing or treating, determining predisposition to diseases and  
CC conditions especially relating to prostate cancer. (I) and its expression  
CC products are used in the diagnostic test to assay for presence of cancer  
CC e.g., in tissue sections, in biopsy sample, in total RNA, in lymph, in  
CC blood etc. (I) is useful for assessing cancer e.g., to determine the type  
CC of cancer, its stage of development, the nature of genetic defect, etc.  
CC The polypeptide encoded by (I) can be used as target for therapy or drug  
CC discovery. (I) can also be used for expressing the polypeptide and thus  
CC for searching specific binding partners of the polypeptide. (I) is useful  
CC in therapeutic applications to treat prostate cancer. The identification  
CC of specific genes, and groups of genes, expressed in pathways  
CC physiologically relevant to prostate cancer permits the definition of  
CC functional and disease pathways and the delineation of targets in these  
CC pathways which are useful in diagnostic, therapeutic, and clinical  
CC applications. This sequence encodes a protein differentially regulated in  
CC prostate cancer  
XX  
SQ Sequence 647 BP; 143 A; 211 C; 145 G; 148 T; 0 U; 0 Other;  
Query Match 97.8%; Score 315.8; DB 8; Length 647;  
Best Local Similarity 99.4%; Pred. No. 2.8e-89;  
Matches 317; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 CGGGGCCAGGCCCTATGCTCCACCGCAAGTGCCTGAAACATCTGGCCCTGATTCYGGG 60  
DB 329 CGGGGCCAGGCCCTATGCTCCACCGCAAGTGCCTGAAACATCTGGCCCTGATTCYGGG 388  
QY 61 CATCCTCATGACCAATGGATTCACCTGTTACTGGTATTCGGCTCTGTGACAGTCTACCA 120  
DB 389 CATCCTCATGACCAATGGATTCACCTGTTACTGGTATTCGGCTCTGTGACAGTCTACCA 448  
QY 121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCTAGCTGCAA 180  
Db 449 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCTAGCTGCAA 508  
QY 181 CTTTGCACCTCCACTGTGCAATGCTGGCCCTGACAGCTGGGGTGTGGCCCTGCCCCCT 240  
Db 509 CTTTGCACCTCCACTGTGCAATGCTGGCCCTGACAGCTGGGGTGTGGCCCTGCCCCCT 568  
QY 241 TGGTCTGCCCCCTAGATACAGCAGTATTATACCCACACACCTGCTTACAGTGTCAATCAAT 300  
Db 569 TGGTCTGCCCCCTAGATACAGCAGTATTATACCCACACACCTGCTTACAGTGTCAATCAAT 628  
QY 301 AAAGTGCACGTGCTTGTGA 319  
Db 629 AAAGTGCACGTGCTTGTGA 647  
RESULT 15  
ADD18965  
ID ADD18965 standard; DNA; 647 BP.  
XX  
AC ADD18965;  
XX  
DT 15-JAN-2004 (first entry)  
DE Human disease related protein DNA sequence SeqID454.  
XX  
XX human; disease state; cytostatic; antiinflammatory; ophthalmological;  
KW antiarteriosclerotic; vulnary; gene therapy;  
KW hypoxia-regulated condition; tumorigenesis; angiogenesis; apoptosis;  
KW inflammation; erythropoiesis; glycolysis; gluconeogenesis;  
KW glucose transportation; catecholamine synthesis; iron transport;  
KW nitric oxide synthesis; cancer; ischaemic condition; reperfusion injury;  
KW retinopathy; neonatal stress; pre-eclampsia; atherosclerosis;  
KW inflammatory condition; wound healing; gene; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO2003018621-A2.  
XX  
PD 06-MAR-2003.  
XX  
PF 23-AUG-2002; 2002WO-GB003892.  
PR 23-AUG-2001; 2001GB-00020558.  
PR 05-OCT-2001; 2001GB-00024037.  
XX  
XX (OXFO-) OXFORD BIOMEDICA UK LTD.  
XX  
XX Kingsman SM, White J, Ward NR, Harris RA, Naylor S, Mundy CR;  
DR WPI; 2003-230046/28.  
DR P-PSDB; ADD18964.  
XX  
XX New substantially purified polypeptide, useful for diagnosing or treating  
PT a hypoxia-regulated condition, such as cancer, ischemia, reperfusion  
PT injury, retinopathy, pre-eclampsia, atherosclerosis, inflammation, or  
PT wound healing.  
XX  
PS Claim 27; SEQ ID NO 454; 424pp; English.  
XX  
XX This invention relates to novel human genes and gene product which are  
CC implicated in certain disease states. Compounds which modulate the  
CC proteins of the invention may have cytostatic, antiinflammatory, the  
CC ophthalmological, antiarteriosclerotic or vulnary activities. The  
CC sequences of the invention may be useful for gene therapy. The invention  
CC may be useful for diagnosing or treating a hypoxia-regulated condition,  
CC such as tumorigenesis, angiogenesis, apoptosis, inflammation,  
CC erythropoiesis, or the biological response to hypoxia conditions  
CC including processes such as glycolysis, gluconeogenesis, glucose  
CC transportation, catecholamine synthesis, iron transport or nitric oxide  
CC synthesis. The disease includes cancer, ischaemic conditions, reperfusion  
CC injury, retinopathy, neonatal stress, pre-eclampsia, atherosclerosis,  
CC inflammatory conditions or wound healing. The present sequence is that of



```
CC a disease related protein encoding DNA sequence of the invention.
XX
SQ Sequence 647 BP; 143 A; 211 C; 145 G; 148 T; 0 U; 0 Other;

Query Match 97.8%; Score 315.8; DB 10; Length 647;
Best Local Similarity 99.4%; Pred. No. 2.8e-89;
Matches 317; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGGGCCAGGCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTTCTGGG 60
Db |||||||
Qy 329 CGGGGCCAGGCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTTCTGGG 388
Db |||||||
Qy 61 CATCCTCATGACCATTTGGATTACCCCTGTTACTGGTATTTCGGCTCTGTGACAGTCTACCA 120
Db |||||||
Qy 389 CATCCTCATGACCATTTGGATTACCCCTGTTACTGGTATTTCGGCTCTGTGACAGTCTACCA 448
Db |||||||
Qy 121 TATTATGTTACAGATAATACAGGAACACGGGGTTACTAGTAGCGCCCATAGCCTGCAA 180
Db |||||||
Qy 449 TATTATGTTACAGATAATACAGGAACACGGGGTTACTAGTAGCGCCCATAGCCTGCAA 508
Db |||||||
Qy 181 CCTTTGCACTCCACTGTGCAATGCTTGGCCCTGCAAGCTGGGGCTGTGGCCCTGCCCCCT 240
Db |||||||
Qy 509 CCTTTGCACTCCACTGTGCAATGCTTGGCCCTGCAAGCTGGGGCTGTGGCCCTGCCCCCT 568
Db |||||||
Qy 241 TGGTCCTGCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATTCAAT 300
Db |||||||
Qy 569 TGGTCCTGCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATTCAAT 628
Db |||||||
Qy 301 AAAGTGCACGTGCTTGTGA 319
Db |||||||
Qy 629 AAAGTGCACGTGCTTGTGA 647
```

Search completed: December 7, 2005, 00:03:07  
Job time : 510.782 secs

*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 11:31:40 ; Search time 804.279 Seconds  
(without alignments)  
3320.997 Million cell updates/sec

Title: US-09-980-046B-6

Perfect score: 323

Sequence: 1 cggggccagcctatgcct.....gtgcacgtgttggaataa 323

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA Main:

1: /cgn2\_6/ptodata/1/pubna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubna/US08\_PUBCOMB.seq.\*  
3: /cgn2\_6/ptodata/1/pubna/US09A\_PUBCOMB.seq.\*  
4: /cgn2\_6/ptodata/1/pubna/US09B\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubna/US10A\_PUBCOMB.seq.\*  
6: /cgn2\_6/ptodata/1/pubna/US10B\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubna/US10C\_PUBCOMB.seq.\*  
8: /cgn2\_6/ptodata/1/pubna/US10D\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubna/US10E\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubna/US11\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	ID	Description
1	319.8	99.0	1421	US-09-814-353-19857
2	319.4	98.9	1095	Sequence 19857, A
3	318.2	98.5	593	Sequence 263, App
4	317.4	98.3	1617	Sequence 3967, Ap
5	317.4	98.3	21478	Sequence 89, Appl
6	315.8	97.8	647	Sequence 201, App
7	315.8	97.8	647	Sequence 484, Ap
8	315.8	97.8	647	Sequence 80, Appl
9	315.8	97.8	647	Sequence 25, Appl
10	315.8	97.8	647	Sequence 934, App
11	315.8	97.8	647	Sequence 35, Appl
12	315.8	97.8	647	Sequence 229, App
13	315.8	97.8	647	Sequence 76, Appl
14	315.8	97.8	647	Sequence 1743, Ap
15	315.8	97.8	842	Sequence 1743, Ap
16	315.8	97.8	842	Sequence 1743, Ap
17	315.8	97.8	842	Sequence 1743, Ap
18	315.8	97.8	842	Sequence 1743, Ap
19	315.8	97.8	853	Sequence 2108, Ap
20	315.8	97.8	853	Sequence 626, App
21	315.8	97.8	853	Sequence 392, App
22	315.8	97.8	853	Sequence 195, App
23	315.8	97.8	853	Sequence 57, Appl

## ALIGNMENTS

## RESULT 1

US-09-814-353-19857  
; Sequence 19857, Application US/09814353  
; Publication No. US20030165831A1  
; GENERAL INFORMATION:  
; APPLICANT: Lee, John  
; APPLICANT: Thompson, Pamela  
; APPLICANT: Lillie, James  
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR  
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER  
; FILE REFERENCE: MRI-006B  
; CURRENT APPLICATION NUMBER: US/09/814,353  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/191,031  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: US 60/207,124  
; PRIOR FILING DATE: 2000-05-25  
; PRIOR APPLICATION NUMBER: US 60/211,940  
; PRIOR FILING DATE: 2000-06-15  
; PRIOR APPLICATION NUMBER: US 60/216,820  
; PRIOR FILING DATE: 2000-07-07  
; PRIOR APPLICATION NUMBER: US 60/220,661  
; PRIOR FILING DATE: 2000-07-25  
; PRIOR APPLICATION NUMBER: US 60/257,672  
; PRIOR FILING DATE: 2000-12-21  
; NUMBER OF SEQ ID NOS: 22037  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 19857  
; LENGTH: 1421  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-814-353-19857

Query Match 99.0%; Score 319.8; DB 3; Length 1421;  
Best Local Similarity 99.4%; Pred. No. 1.7e-98;  
Matches 321; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGGGCCCGAGCCCTATGCTCCACCGCAAGTGGTGAACATCTGGGCCCTGATTCGGG 60  
DB 420 CGGGCCCGAGCCCTATGCTCCACCGCAAGTGGTGAACATCTGGGCCCTGATTCGGG 479  
QY 61 CATCTCTATGACCAATGATTCACCTCTTACTGTTATCGGCTCTGTGACAGTCTACCA 120  
DB 480 CATCTCTATGACCAATGATTCACCTCTTACTGTTATCGGCTCTGTGACAGTCTACCA 539

Sequence 10, Appl  
Sequence 52, Appl  
Sequence 4116, Ap  
Sequence 283, App  
Sequence 283, App  
Sequence 283, App  
Sequence 283, App  
Sequence 4078, Ap  
Sequence 11579, A  
Sequence 3, Appli  
Sequence 5, Appli  
Sequence 27, Appl  
Sequence 14867, A  
Sequence 53, Appl  
Sequence 4224, Ap  
Sequence 116, App  
Sequence 4364, Ap  
Sequence 14920, A  
Sequence 35636, A  
Sequence 4345, Ap

Qy 121 TATTATGTTACAGATAATACAGAAAAACGGGGTTACTAGTAGCCGCCATAGCCTGCAA 180  
Db |||||  
Qy 540 TATTATGTTACAGATAATACAGAAAAACGGGGTTACTAGTAGCCGCCATAGCCTGCAA 599  
Db |||||  
Qy 181 CTTTGGCACTCCACTGTGCAATGTGCGCTTGCACGCTGGGGCTGTGGCCCTGCCCCCT 240  
Db CTTTGGCACTCCACTGTGCAATGTGCGCTTGCACGCTGGGGCTGTGGCCCTGCCCCCT 659  
Qy 241 TGGTCCCTGCCCTAGATACAGAGTTTATATACCCACACACCTGTCTACAGTGTCAATCAAT 300  
Db |||||  
Qy 660 TGGTCCCTGCCCTAGATACAGAGTTTATATACCCACACACCTGTCTACAGTGTCAATCAAT 719  
Db |||||  
Qy 301 AAAGTGACAGTGTGTGAAAAA 323  
Db |||||  
Qy 720 AAAGTGACAGTGTGTGAAAAA 742  
Db |||||  
RESULT 2  
US-10-240-965-263  
; Sequence 263, Application US/10240965  
; Publication No. US20030165924A1  
; GENERAL INFORMATION:  
; APPLICANT: INCYTE GENOMICS, INC.  
; APPLICANT: SHIFFMAN, Dov  
; APPLICANT: SOMOGYI, Roland  
; APPLICANT: LAWN, Richard M.  
; APPLICANT: SEILHAMER, Jeffrey J.  
; APPLICANT: PORTER, Gordon J.  
; APPLICANT: MIKITA, Thomas  
; APPLICANT: TAL, Julie  
; TITLE OF INVENTION: GENES EXPRESSED IN FOAM CELL DIFFERENTIATION  
; FILE REFERENCE: PA-0025 PCT  
; CURRENT APPLICATION NUMBER: US/10/240,965  
; CURRENT FILING DATE: 2002-10-04  
; PRIOR APPLICATION NUMBER: 60/195,106  
; PRIOR FILING DATE: 2000-04-05  
; NUMBER OF SEQ ID NOS: 276  
; SOFTWARE: PERL Program  
; SEQ ID NO 263  
; LENGTH: 1095  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. US20030165924A1 988231.7  
US-10-240-965-263  
Query Match 98.9%; Score 319.4; DB 6; Length 1095;  
Best Local Similarity 99.7%; Pred. No. 2.le-98;  
Matches 320; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 CCGGGCCAGCCCTATGCTCCACCGCAAGTGCCTGAAACATCTGGGCCCTGATTTGGG 60  
Db 775 CCGGGCCAGCCCTATGCTCCACCGCAAGTGCCTGAAACATCTGGGCCCTGATTTGGG 834  
Qy 61 CATCCTCATGACCATTTGATTCACCCCTGTTACTGGTATTTCGGCTCTGTGACAGTCTACCA 120  
Db 835 CATCCTCATGACCATTTGATTCACCCCTGTTACTGGTATTTCGGCTCTGTGACAGTCTACCA 894  
Qy 121 TATTATGTTACAGATAATACAGAAAAACGGGGTTACTAGTAGCCGCCATAGCCTGCAA 180  
Db 895 TATTATGTTACAGATAATACAGAAAAACGGGGTTACTAGTAGCCGCCATAGCCTGCAA 954  
Qy 181 CTTTGGCACTCCACTGTGCAATGTGCGCTTGCACGCTGGGGCTGTGGCCCTGCCCCCT 240  
Db 955 CTTTGGCACTCCACTGTGCAATGTGCGCTTGCACGCTGGGGCTGTGGCCCTGCCCCCT 1014  
Qy 241 TGGTCCCTGCCCTAGATACAGAGTTTATATACCCACACACCTGTCTACAGTGTCAATCAAT 300  
Db 1015 TGGTCCCTGCCCTAGATACAGAGTTTATATACCCACACACCTGTCTACAGTGTCAATCAAT 1074  
Qy 301 AAAGTGACAGTGTGTGAAAA 321  
Db |||||

Db 1075 AAAGTGACAGTGTGTGAAAA 1095  
RESULT 3  
US-09-969-034-3967/c  
; Sequence 3967, Application US/09969034  
; Publication No. US20040110668A1  
; GENERAL INFORMATION:  
; APPLICANT: Burgess, Christopher C.  
; APPLICANT: Aetle, Jon H.  
; APPLICANT: Carroll, Eddie III  
; APPLICANT: Catino, Theodore J.  
; APPLICANT: Dwivedi, Poornima  
; APPLICANT: Molino, Gary A.  
; APPLICANT: Thiagalingam, Arunthathi  
; APPLICANT: Lewis, Marcia E.  
; TITLE OF INVENTION: Nucleic Acid Sequences Differentially  
; FILE REFERENCE: 1657/1032  
; CURRENT APPLICATION NUMBER: US/09/969,034  
; CURRENT FILING DATE: 2001-10-02  
; PRIOR APPLICATION NUMBER: 60/237,271  
; PRIOR FILING DATE: 2000-02-10  
; NUMBER OF SEQ ID NOS: 4494  
; SOFTWARE: Fast-SEQ for Windows Version 4.0  
; SEQ ID NO 3967  
; LENGTH: 593  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 380, 387, 407, 415, 424, 433, 436, 442, 467, 472, 493, 496,  
; LOCATION: 520, 524, 533, 545, 548, 556, 559, 563  
; OTHER INFORMATION: n = A,T,C or G  
US-09-969-034-3967  
Query Match 98.5%; Score 318.2; DB 3; Length 593;  
Best Local Similarity 99.1%; Pred. No. 4.le-98;  
Matches 320; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
Qy 1 CCGGGCCAGCCCTATGCTCCACCGCCAAAGTGCCTGAAACATCTGGGCCCTGATTTGGG 60  
Db 342 CCGGGCCAGCCCTATGCTCCACCGCCAAAGTGCCTGAAACATCTGGGCCCTGATTTGGG 283  
Qy 61 CATCCTCATGACCATTTGATTCACCCCTGTTACTGGTATTTCGGCTCTGTGACAGTCTACCA 120  
Db 282 CATCCTCATGACCATTTGATTCACCCCTGTTACTGGTATTTCGGCTCTGTGACAGTCTACCA 223  
Qy 121 TATTATGTTACAGATAATACAGAAAAACGGGGTTACTAGTAGCCGCCATAGCCTGCAA 180  
Db 222 TATTATGTTACAGATAATACAGAAAAACGGGGTTACTAGTAGCCGCCATAGCCTGCAA 163  
Qy 181 CTTTGGCACTCCACTGTGCAATGTGCGCTTGCACGCTGGGGCTGTGGCCCTGCCCCCT 240  
Db 162 CTTTGGCACTCCACTGTGCAATGTGCGCTTGCACGCTGGGGCTGTGGCCCTGCCCCCT 103  
Qy 241 TGGTCCCTGCCCTAGATACAGAGTTTATATACCCACACACCTGTCTACAGTGTCAATCAAT 300  
Db 102 TGGTCCCTGCCCTAGATACAGAGTTTATATACCCACACACCTGTCTACAGTGTCAATCAAT 43  
Qy 301 AAAGTGACAGTGTGTGAAAAA 323  
Db 42 AAAGTGACAGTGTGTGAAAAA 20  
RESULT 4  
US-10-247-671-89  
; Sequence 89, Application US/10247671  
; Publication No. US20030194721A1  
; GENERAL INFORMATION:  
; APPLICANT: Mikita, Thomas  
; APPLICANT: Shiffman, Dov  
; APPLICANT: Porter, Gordon, J.

APPLICANT: Kaser, Matthew R.  
TITLE OF INVENTION: GENES EXPRESSED IN TREATED FOAM CELLS  
FILE REFERENCE: PA-0050 US  
CURRENT APPLICATION NUMBER: US/10/247,671  
CURRENT FILING DATE: 2002-09-18  
PRIOR APPLICATION NUMBER: 60/323,784  
PRIOR FILING DATE: 2001-09-19  
NUMBER OF SEQ ID NOS: 186  
SOFTWARE: PERL Program  
SEQ ID NO 89  
LENGTH: 1617  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
OTHER INFORMATION: Incyte ID No. US20030194721A1 1328372.14  
US-10-247-671-89

Query Match 98.3%; Score 317.4; DB 6; Length 1617;  
Best Local Similarity 99.7%; Pred. No. 1.2e-97;  
Matches 318; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CGGGCCAGGCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCCTGGG 60  
DB 773 CGGGCCAGGCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCCTGGG 832  
QY 61 CATCTCATGACCATTTGGATTCCACCTCTGTACTGTGTTATTCGGCTCTGTGACAGTCTACCA 120  
DB 833 CATCTCATGACCATTTGGATTCCCTGTGTACTGTGTTATTCGGCTCTGTGACAGTCTACCA 892  
QY 121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTTACTAGTAGCGGCCCATAGCCTGCAA 180  
DB 893 TATTATGTTACAGATAATACAGGAAAAACGGGGTTTACTAGTAGCGGCCCATAGCCTGCAA 952  
QY 181 CTTTGGCACTCCACTGTGCAATGCTGGCCCTGCACGCTGGGGCTGTGGCCCTGCCCCCT 240  
DB 953 CTTTGGCACTCCACTGTGCAATGCTGGCCCTGCACGCTGGGGCTGTGGCCCTGCCCCCT 1012  
QY 241 TGGTCTGCCCTAGATACAGAGTTTATACCCACACACCTGTCTACAGTGTCTATCAAT 300  
DB 1013 TGGTCTGCCCTAGATACAGAGTTTATACCCACACAGTTTATACCCACACCTGTCTACAGTGTCTATCAAT 1072  
QY 301 AAAGTGCACGTCTGTGTA 319  
DB 1073 AAAGTGCACGTCTGTGTA 1091

## RESULT 5

US-10-367-094-201  
Sequence 201, Application US/10367094  
Publication No. US20040170982A1  
GENERAL INFORMATION:  
APPLICANT: David W. Morris  
APPLICANT: Marc Malandro  
TITLE OF INVENTION: Novel Therapeutic Targets in Cancer  
FILE REFERENCE: 529452001500  
CURRENT APPLICATION NUMBER: US/10/367,094  
CURRENT FILING DATE: 2003-02-14  
NUMBER OF SEQ ID NOS: 203  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 201  
LENGTH: 21478  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1)-(21478)  
OTHER INFORMATION: n = A,T,C or G  
US-10-367-094-201

Query Match 98.3%; Score 317.4; DB 7; Length 21478;  
Best Local Similarity 99.7%; Pred. No. 3.9e-97;  
Matches 318; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGGCCAGGCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCCTGGG 60  
DB 17039 CGGGCCAGGCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCCTGGG 17098  
QY 61 CATCTCATGACCATTTGGATTCCACCTGTGTACTGTGTTATTCGGCTCTGTGACAGTCTACCA 120  
DB 17099 CATCTCATGACCATTTGGATTCCCTGTGTACTGTGTTATTCGGCTCTGTGACAGTCTACCA 17158  
QY 121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTTACTAGTAGCGGCCCATAGCCTGCAA 180  
DB 17159 TATTATGTTACAGATAATACAGGAAAAACGGGGTTTACTAGTAGCGGCCCATAGCCTGCAA 17218  
QY 181 CTTTGGCACTCCACTGTGCAATGCTGGCCCTGCACGCTGGGGCTGTGGCCCTGCCCCCT 240  
DB 17219 CTTTGGCACTCCACTGTGCAATGCTGGCCCTGCACGCTGGGGCTGTGGCCCTGCCCCCT 17278  
QY 241 TGGTCTGCCCTAGATACAGAGTTTATACCCACACACCTGTCTACAGTGTCTATCAAT 300  
DB 17279 TGGTCTGCCCTAGATACAGAGTTTATACCCACACACCTGTCTACAGTGTCTATCAAT 17338  
QY 301 AAAGTGCACGTCTGTGTA 319  
DB 17339 AAAGTGCACGTCTGTGTA 17357

## RESULT 6

US-09-969-034-4484  
Sequence 4484, Application US/09969034  
Publication No. US20040110668A1  
GENERAL INFORMATION:  
APPLICANT: Burgess, Christopher C.  
APPLICANT: Astle, Jon H.  
APPLICANT: Carroll, Eddie III  
APPLICANT: Catino, Theodore J.  
APPLICANT: Dwivedi, Poorima  
APPLICANT: Molino, Gary A.  
APPLICANT: Thiagalingam, Arunthathi  
APPLICANT: Lewis, Marcia E.  
TITLE OF INVENTION: Nucleic Acid Sequences Differentially  
Expressed in Cancer Tissue  
FILE REFERENCE: 1657/1032  
CURRENT APPLICATION NUMBER: US/09/969,034  
CURRENT FILING DATE: 2001-10-02  
PRIOR APPLICATION NUMBER: 60/237,271  
PRIOR FILING DATE: 2000-02-10  
NUMBER OF SEQ ID NOS: 4494  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 4484  
LENGTH: 647  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-969-034-4484

Query Match 97.8%; Score 315.8; DB 3; Length 647;  
Best Local Similarity 99.4%; Pred. No. 2.8e-97;  
Matches 317; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 CGGGCCAGGCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCCTGGG 60  
DB 329 CGGGCCAGGCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCCTGGG 388  
QY 61 CATCTCATGACCATTTGGATTCCACCTGTGTACTGTGTTATTCGGCTCTGTGACAGTCTACCA 120  
DB 389 CATCTCATGACCATTTGGATTCCCTGTGTACTGTGTTATTCGGCTCTGTGACAGTCTACCA 448  
QY 121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTTACTAGTAGCGGCCCATAGCCTGCAA 180  
DB 449 TATTATGTTACAGATAATACAGGAAAAACGGGGTTTACTAGTAGCGGCCCATAGCCTGCAA 508  
QY 181 CTTTGGCACTCCACTGTGCAATGCTGGCCCTGCACGCTGGGGCTGTGGCCCTGCCCCCT 240  
DB 509 CTTTGGCACTCCACTGTGCAATGCTGGCCCTGCACGCTGGGGCTGTGGCCCTGCCCCCT 568

Qy 241 TGGTCTGCCCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 300  
Db 569 TGGTCTGCCCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 628  
Qy 301 AAAGTGCACGTGCTTGGA 319  
Db 629 AAAGTGCACGTGCTTGGA 647

RESULT 7  
US-10-301-822-80  
; Sequence 80, Application US/10301822  
; Publication No. US20030148410A1  
; GENERAL INFORMATION:  
; APPLICANT: Millennium Pharmaceuticals, Inc.  
; APPLICANT: Berger, Allison  
; APPLICANT: Guillemette, Tracy L.  
; APPLICANT: Kamatkar, Shubhangi  
; APPLICANT: Schlegel, Robert  
; APPLICANT: Monahan, John E.  
; APPLICANT: Thibodeau, Stephen N.  
; APPLICANT: Burgart, Lawrence J.  
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND  
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
; FILE OF INVENTION: THERAPY OF COLON CANCER  
; FILE REFERENCE: MP01-029P2RNM  
; CURRENT APPLICATION NUMBER: US/10/301,822  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US 60/339,971  
; PRIOR FILING DATE: 2001-12-10  
; PRIOR APPLICATION NUMBER: US 60/361,978  
; PRIOR FILING DATE: 2002-03-05  
; PRIOR APPLICATION NUMBER: US 60/381,988  
; PRIOR FILING DATE: 2002-05-20  
; NUMBER OF SEQ ID NOS: 228  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 80  
; LENGTH: 647  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (111)...(488)  
US-10-301-822-80

Query Match 97.8%; Score 315.8; DB 6; Length 647;  
Best Local Similarity 99.4%; Pred. No. 2.8e-97;  
Matches 317; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGGGCCAGCCCTATCGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATCTGGG 60  
Db 329 CGGGGCCAGCCCTATCGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATCTGGG 388  
Qy 61 CATCCTCATGACCAATTGGATTACCCCTGTTACTGGTATTGGCTCTGTGACGTCTACCA 120  
Db 389 CATCCTCATGACCAATTGGATTACCCCTGTTACTGGTATTGGCTCTGTGACGTCTACCA 448  
Qy 121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTACGCCGCCATAGCCTGCAA 180  
Db 449 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTACGCCGCCATAGCCTGCAA 508  
Qy 181 CCTTTGCACCTCCACTGTGCAATCTGGCCCTGCACGCTGGGGCTGTTGCCCTGCCCCCT 240  
Db 509 CCTTTGCACCTCCACTGTGCAATCTGGCCCTGCACGCTGGGGCTGTTGCCCTGCCCCCT 568  
Qy 241 TGGTCTGCCCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 300  
Db 569 TGGTCTGCCCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 628  
Qy 301 AAAGTGCACGTGCTTGGA 319  
Db 629 AAAGTGCACGTGCTTGGA 647

RESULT 9  
US-10-172-118-934  
; Sequence 934, Application US/10172118  
; Publication No. US20030224374A1  
; GENERAL INFORMATION:  
; APPLICANT: Dai, Hongyue  
; APPLICANT: He, Yudong  
; APPLICANT: Linsley, Peter  
; APPLICANT: Mao, Mao  
; APPLICANT: Roberts, Chris  
; APPLICANT: Van 't Veijer, Laura  
; APPLICANT: Van de Vijver, Marc  
; APPLICANT: Bernards, Rene  
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients  
; FILE REFERENCE: 9301-175-999  
; CURRENT APPLICATION NUMBER: US/10/172,118  
; CURRENT FILING DATE: 2002-06-14  
; PRIOR APPLICATION NUMBER: 60/380,770  
; PRIOR FILING DATE: 2002-05-14

RESULT 8  
US-10-133-937-25  
; Sequence 25, Application US/10133937  
; Publication No. US20030207278A1  
; GENERAL INFORMATION:  
; APPLICANT: Khan, Javed  
; APPLICANT: Ringner, Markus  
; APPLICANT: Peterson, Carsten  
; APPLICANT: Meltzer, Paul  
; TITLE OF INVENTION: METHODS FOR ANALYZING HIGH DIMENSIONAL DATA FOR CLASSIFYING,  
; TITLE OF INVENTION: DIAGNOSING, PROGNOSTICATING, AND/OR PREDICTING DISEASES AND  
; TITLE OF INVENTION: OTHER BIOLOGICAL STATES  
; FILE REFERENCE: 11613.56US01  
; CURRENT APPLICATION NUMBER: US/10/133,937  
; CURRENT FILING DATE: 2002-11-04  
; NUMBER OF SEQ ID NOS: 99  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 25  
; LENGTH: 647  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-133-937-25

Query Match 97.8%; Score 315.8; DB 6; Length 647;  
Best Local Similarity 99.4%; Pred. No. 2.8e-97;  
Matches 317; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGGGCCAGCCCTATCGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATCTGGG 60  
Db 329 CGGGGCCAGCCCTATCGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATCTGGG 388  
Qy 61 CATCCTCATGACCAATTGGATTACCCCTGTTACTGGTATTGGCTCTGTGACGTCTACCA 120  
Db 389 CATCCTCATGACCAATTGGATTACCCCTGTTACTGGTATTGGCTCTGTGACGTCTACCA 448  
Qy 121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTACGCCGCCATAGCCTGCAA 180  
Db 449 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTACGCCGCCATAGCCTGCAA 508  
Qy 181 CCTTTGCACCTCCACTGTGCAATCTGGCCCTGCACGCTGGGGCTGTTGCCCTGCCCCCT 240  
Db 509 CCTTTGCACCTCCACTGTGCAATCTGGCCCTGCACGCTGGGGCTGTTGCCCTGCCCCCT 568  
Qy 241 TGGTCTGCCCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 300  
Db 569 TGGTCTGCCCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 628  
Qy 301 AAAGTGCACGTGCTTGGA 319  
Db 629 AAAGTGCACGTGCTTGGA 647

RESULT 9  
US-10-172-118-934  
; Sequence 934, Application US/10172118  
; Publication No. US20030224374A1  
; GENERAL INFORMATION:  
; APPLICANT: Dai, Hongyue  
; APPLICANT: He, Yudong  
; APPLICANT: Linsley, Peter  
; APPLICANT: Mao, Mao  
; APPLICANT: Roberts, Chris  
; APPLICANT: Van 't Veijer, Laura  
; APPLICANT: Van de Vijver, Marc  
; APPLICANT: Bernards, Rene  
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients  
; FILE REFERENCE: 9301-175-999  
; CURRENT APPLICATION NUMBER: US/10/172,118  
; CURRENT FILING DATE: 2002-06-14  
; PRIOR APPLICATION NUMBER: 60/380,770  
; PRIOR FILING DATE: 2002-05-14

```
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 934
; LENGTH: 647
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM_003641
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-934

Query Match          97.8%; Score 315.8; DB 6; Length 647;
Best Local Similarity 99.4%; Pred. No. 2.8e-97;
Matches 317; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGGGCCAGGCTATGCTCCACCGCAAGTGCCTGAACTCTGGGCCCTGATTTCTGGG 60
Db 329 CGGGGCCAGGCTATGCTCCACCGCAAGTGCCTGAACTCTGGGCCCTGATTTCTGGG 388

Qy 61 CATCTCATGACCAATGGATTCACTGTTACTGGTATTTCGGCTCTGTGACAGTCTACCA 120
Db 389 CATCTCATGACCAATGGATTCACTGTTACTGGTATTTCGGCTCTGTGAAACAGTCTACCA 448

Qy 121 TATTATGTTACAGATAATACAGGAAACCGGGGTTACTAGTAGCGGCCCATAGCTGCAA 180
Db 389 CATCTCATGACCAATGGATTCACTGTTACTGGTATTTCGGCTCTGTGAAACAGTCTACCA 448

Qy 449 TATTATGTTACAGATAATACAGGAAACCGGGGTTACTAGTAGCGGCCCATAGCTGCAA 508
Db 449 TATTATGTTACAGATAATACAGGAAACCGGGGTTACTAGTAGCGGCCCATAGCTGCAA 508

Qy 181 CTTTGGCACTCCACTGTCGAATGCTGGCCCTGCAAGCTGGGGCTGTTGGCCCTGCCCCCT 240
Db 509 CTTTGGCACTCCACTGTCGAATGCTGGCCCTGCAAGCTGGGGCTGTTGGCCCTGCCCCCT 568

Qy 241 TGGTCTGCTGCCCTAGATACAGAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 300
Db 569 TGGTCTGCTGCCCTAGATACAGAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 628

Qy 301 AAAGTGCAGCTGCTTGTA 319
Db 629 AAAGTGCAGCTGCTTGTA 647

RESULT 10
US-10-159-563-25
; Sequence 25, Application US/10159563
; Publication No. US20040009154A1
; GENERAL INFORMATION:
; APPLICANT: Khan, Javed
; APPLICANT: Ringner, Markus
; APPLICANT: Peterson, Carsten
; APPLICANT: Meltzer, Paul
; TITLE OF INVENTION: SELECTIONS OF GENES AND METHODS OF USING THE SAME FOR
; FILE REFERENCE: 11613.56US11
; CURRENT APPLICATION NUMBER: US/10/159,563
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: US 10/133,937
; PRIOR FILING DATE: 2002-04-25
; NUMBER OF SEQ ID NOS: 444
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 25
; LENGTH: 647
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-159-563-25

Query Match          97.8%; Score 315.8; DB 6; Length 647;
Best Local Similarity 99.4%; Pred. No. 2.8e-97;
Matches 317; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGGGCCAGGCTATGCTCCACCGCAAGTGCCTGAACTCTGGGCCCTGATTTCTGGG 60
Db 329 CGGGGCCAGGCTATGCTCCACCGCAAGTGCCTGAACTCTGGGCCCTGATTTCTGGG 388

Qy 61 CATCTCATGACCAATGGATTCACTGTTACTGGTATTTCGGCTCTGTGACAGTCTACCA 120
Db 389 CATCTCATGACCAATGGATTCACTGTTACTGGTATTTCGGCTCTGTGAAACAGTCTACCA 448

Qy 121 TATTATGTTACAGATAATACAGGAAACCGGGGTTACTAGTAGCGGCCCATAGCTGCAA 180
Db 389 CATCTCATGACCAATGGATTCACTGTTACTGGTATTTCGGCTCTGTGAAACAGTCTACCA 448

Qy 449 TATTATGTTACAGATAATACAGGAAACCGGGGTTACTAGTAGCGGCCCATAGCTGCAA 508
Db 449 TATTATGTTACAGATAATACAGGAAACCGGGGTTACTAGTAGCGGCCCATAGCTGCAA 508

Qy 181 CTTTGGCACTCCACTGTCGAATGCTGGCCCTGCAAGCTGGGGCTGTTGGCCCTGCCCCCT 240
Db 509 CTTTGGCACTCCACTGTCGAATGCTGGCCCTGCAAGCTGGGGCTGTTGGCCCTGCCCCCT 568

Qy 241 TGGTCTGCTGCCCTAGATACAGAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 300
Db 569 TGGTCTGCTGCCCTAGATACAGAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 628

Qy 301 AAAGTGCAGCTGCTTGTA 319
Db 629 AAAGTGCAGCTGCTTGTA 647

Query Match          97.8%; Score 315.8; DB 7; Length 647;
Best Local Similarity 99.4%; Pred. No. 2.8e-97;
Matches 317; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGGGCCAGGCTATGCTCCACCGCAAGTGCCTGAACTCTGGGCCCTGATTTCTGGG 60
Db 329 CGGGGCCAGGCTATGCTCCACCGCAAGTGCCTGAACTCTGGGCCCTGATTTCTGGG 388

Qy 61 CATCTCATGACCAATGGATTCACTGTTACTGGTATTTCGGCTCTGTGACAGTCTACCA 120
Db 389 CATCTCATGACCAATGGATTCACTGTTACTGGTATTTCGGCTCTGTGAAACAGTCTACCA 448

Qy 121 TATTATGTTACAGATAATACAGGAAACCGGGGTTACTAGTAGCGGCCCATAGCTGCAA 180
Db 449 TATTATGTTACAGATAATACAGGAAACCGGGGTTACTAGTAGCGGCCCATAGCTGCAA 508

Qy 181 CTTTGGCACTCCACTGTCGAATGCTGGCCCTGCAAGCTGGGGCTGTTGGCCCTGCCCCCT 240
Db 509 CTTTGGCACTCCACTGTCGAATGCTGGCCCTGCAAGCTGGGGCTGTTGGCCCTGCCCCCT 568

Qy 241 TGGTCTGCTGCCCTAGATACAGAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 300
Db 569 TGGTCTGCTGCCCTAGATACAGAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 628
```

Qy 301 AAAGTGCACGTGCTTGGA 319  
|  
Db 629 AAAGTGCACGTGCTTGGA 647

## RESULT 12

US-10-734-564-35  
; Sequence 35, Application US/10734564  
; Publication No. US20040157278A1  
; GENERAL INFORMATION:  
; APPLICANT: Christopher C Burgess et al  
; TITLE OF INVENTION: Detection Methods Using TIMP1  
; FILE REFERENCE: 1657/2012  
; CURRENT APPLICATION NUMBER: US/10/734,564  
; CURRENT FILING DATE: 2003-12-12  
; NUMBER OF SEQ ID NOS: 138  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 35  
; LENGTH: 647  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-734-564-35

Query Match 97.8%; Score 315.8; DB 7; Length 647;  
Best Local Similarity 99.4%; Pred. No. 2.8e-97;  
Matches 317; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGGGCCAGCCCTATGCTCCACCGCAAGTGCTGAAACATCTGGGCCCTGATTCGGG 60  
Db 329 CGGGGCCAGCCCTATGCTCCACCGCAAGTGCTGAAACATCTGGGCCCTGATTCGGG 388  
  
Qy 61 CATCCTCATGACCAATGGATTACCCCTGTTACTCGTATTGGCTCTGTGACAGTCTACCA 120  
Db 389 CATCCTCATGACCAATGGATTACCCCTGTTACTCGTATTGGCTCTGTGACAGTCTACCA 448  
  
Qy 121 TATTATGTTACAGATAATACAGAAAAACGGGGTTACTAGTACGCCGCCCATAGCTGCAA 180  
Db 449 TATTATGTTACAGATAATACAGAAAAACGGGGTTACTAGTACGCCGCCCATAGCTGCAA 508  
  
Qy 181 CCTTGTCACTCCACTGTGCAATGCTGGCCCTGCACGCTGGGGCTGTTGCCCTGCCCCCT 240  
Db 509 CCTTGTCACTCCACTGTGCAATGCTGGCCCTGCACGCTGGGGCTGTTGCCCTGCCCCCT 568  
  
Qy 241 TGGTCTCTGCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 300  
Db 569 TGGTCTCTGCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 628  
  
Qy 301 AAAGTGCACGTGCTTGGA 319  
|  
Db 629 AAAGTGCACGTGCTTGGA 647

## RESULT 13

US-10-755-889-229  
; Sequence 229, Application US/10755889  
; Publication No. US20040171823A1  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-KB  
; FILE REFERENCE: D0284 NP  
; CURRENT APPLICATION NUMBER: US/10/755,889  
; CURRENT FILING DATE: 2004-01-13  
; PRIOR APPLICATION NUMBER: U.S. 60/440,068  
; PRIOR FILING DATE: 2003-01-14  
; PRIOR APPLICATION NUMBER: U.S. 60/469,757  
; PRIOR FILING DATE: 2003-05-12  
; NUMBER OF SEQ ID NOS: 823  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 229  
; LENGTH: 647  
; TYPE: DNA  
; ORGANISM: Homo sapiens

US-10-755-889-229  
Query Match 97.8%; Score 315.8; DB 7; Length 647;  
Best Local Similarity 99.4%; Pred. No. 2.8e-97;  
Matches 317; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
Qy 1 CGGGGCCAGCCCTATGCTCCACCGCAAGTGCTGAAACATCTGGGCCCTGATTCGGG 60  
Db 329 CGGGGCCAGCCCTATGCTCCACCGCAAGTGCTGAAACATCTGGGCCCTGATTCGGG 388  
  
Qy 61 CATCCTCATGACCAATGGATTACCCCTGTTACTCGTATTGGCTCTGTGACAGTCTACCA 120  
Db 389 CATCCTCATGACCAATGGATTACCCCTGTTACTCGTATTGGCTCTGTGACAGTCTACCA 448  
  
Qy 121 TATTATGTTACAGATAATACAGAAAAACGGGGTTACTAGTACGCCGCCCATAGCTGCAA 180  
Db 449 TATTATGTTACAGATAATACAGAAAAACGGGGTTACTAGTACGCCGCCCATAGCTGCAA 508  
  
Qy 181 CCTTGTCACTCCACTGTGCAATGCTGGCCCTGCACGCTGGGGCTGTTGCCCTGCCCCCT 240  
Db 509 CCTTGTCACTCCACTGTGCAATGCTGGCCCTGCACGCTGGGGCTGTTGCCCTGCCCCCT 568  
  
Qy 241 TGGTCTCTGCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 300  
Db 569 TGGTCTCTGCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 628  
  
Qy 301 AAAGTGCACGTGCTTGGA 319  
|  
Db 629 AAAGTGCACGTGCTTGGA 647

## RESULT 14

US-10-631-467-76  
; Sequence 76, Application US/10631467  
; Publication No. US20050208496A1  
; GENERAL INFORMATION:  
; APPLICANT: Genex Research Inc.  
; TITLE OF INVENTION: Method for testing for bronchial asthma, or chronic obstructive p  
; FILE REFERENCE: 3462.1005-000  
; CURRENT APPLICATION NUMBER: US/10/631,467  
; CURRENT FILING DATE: 2003-07-31  
; PRIOR APPLICATION NUMBER: JP 2003-077212  
; PRIOR FILING DATE: 2003-03-20  
; PRIOR APPLICATION NUMBER: JP 2002-229312  
; PRIOR FILING DATE: 2002-08-06  
; NUMBER OF SEQ ID NOS: 2086  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 76  
; LENGTH: 647  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-631-467-76

Query Match 97.8%; Score 315.8; DB 9; Length 647;  
Best Local Similarity 99.4%; Pred. No. 2.8e-97;  
Matches 317; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
Qy 1 CGGGGCCAGCCCTATGCTCCACCGCAAGTGCTGAAACATCTGGGCCCTGATTCGGG 60  
Db 329 CGGGGCCAGCCCTATGCTCCACCGCAAGTGCTGAAACATCTGGGCCCTGATTCGGG 388  
  
Qy 61 CATCCTCATGACCAATGGATTACCCCTGTTACTCGTATTGGCTCTGTGACAGTCTACCA 120  
Db 389 CATCCTCATGACCAATGGATTACCCCTGTTACTCGTATTGGCTCTGTGACAGTCTACCA 448  
  
Qy 121 TATTATGTTACAGATAATACAGAAAAACGGGGTTACTAGTACGCCGCCCATAGCTGCAA 180  
Db 449 TATTATGTTACAGATAATACAGAAAAACGGGGTTACTAGTACGCCGCCCATAGCTGCAA 508  
  
Qy 181 CCTTGTCACTCCACTGTGCAATGCTGGCCCTGCACGCTGGGGCTGTTGCCCTGCCCCCT 240  
Db 509 CCTTGTCACTCCACTGTGCAATGCTGGCCCTGCACGCTGGGGCTGTTGCCCTGCCCCCT 568



Qy 241 TGGTCCTGCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 300  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
569 TGGTCCTGCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 628  
Qy 301 AAAGTGCACGTGCTTGGA 319  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
629 AAAGTGCACGTGCTTGGA 647

## RESULT 15

US-09-920-300A-1743  
; Sequence 1743, Application US/09920300A  
; Patent No. US20020136728A1  
; GENERAL INFORMATION:  
; APPLICANT: King, Gordon E.  
; APPLICANT: Mesgher, Madeleine Joy  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Secrist, Heather  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER  
; FILE REFERENCE: 210121.547  
; CURRENT APPLICATION NUMBER: US/09/920.300A  
; CURRENT FILING DATE: 2001-07-31  
; NUMBER OF SEQ ID NOS: 1789  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 1743  
; LENGTH: 842  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-920-300A-1743

Query Match 97.8%; Score 315.8; DB 3; Length 842;  
Best Local Similarity 99.4%; Pred. No. 3.2e-97;  
Matches 317; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 CGGGGCCAGGCTATGCTCCACGGCAAGTGCTGAACATCTGGGCCCTGATTCGGG 60  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
524 CGGGGCCAGGCTATGCTCCACGGCAAGTGCTGAACATCTGGGCCCTGATTCGGG 583  
Qy 61 CATCTCATGACCATTTGGAATTCACCTGTCTTACTGTGTGTTTGGCTCTGTGACAGTCTACCA 120  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
584 CATCTCATGACCATTTGGAATTCATCTGTCTACTGTGTATTCGGCTCTGTGACAGTCTACCA 643  
Qy 121 TATTATGTTACAGATAATACAGGAARACGGGGTTACTAGTAGCGCCCATAGCCTGCAA 180  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
644 TATTATGTTACAGATAATACAGGAARACGGGGTTACTAGTAGCGCCCATAGCCTGCAA 703  
Qy 181 CCTTTGCACTCCACTGTGCAATGCTTGGCCCTGCAAGCTGGGGCTGTGTCGCCCTGCCCT 240  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
704 CCTTTGCACTCCACTGTGCAATGCTTGGCCCTGCAAGCTGGGGCTGTGTCGCCCTGCCCT 763  
Qy 241 TGGTCCTGCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 300  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
764 TGGTCCTGCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 823  
Qy 301 AAAGTGCACGTGCTTGGA 319  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
824 AAAGTGCACGTGCTTGGA 842

Search completed: December 6, 2005, 22:57:07  
Job time : 806.279 secs

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 11:34:11 ; Search time 220.487 Seconds  
(without alignments)  
455.919 Million cell updates/sec

Title: US-09-980-046B-6  
Perfect score: 323  
Sequence: 1 cggggccaggcctatgcct.....gtgcagctgttgtaaaaa 323

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3289935 seqs, 155610033 residues  
Total number of hits satisfying chosen parameters: 6579870

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA New:  
1: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB\_seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB\_seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB\_seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB\_seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB\_seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB\_seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB\_seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB\_seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB\_seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB\_seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	116.2	36.0	182314	7	US-11-112-908-45
C 2	78.4	24.3	402	6	US-10-821-234-582
C 3	71	22.0	3667	6	US-10-750-185-25113
C 4	62.4	19.3	176	7	US-11-108-172-509
C 5	34	10.5	735	6	US-10-750-185-59126
C 6	32.6	10.1	923	6	US-10-750-185-32680
C 7	31	9.6	3085	6	US-10-750-185-51883
C 8	30.4	9.4	1648	6	US-10-750-185-55588
C 9	30	9.3	1242	6	US-10-750-185-30919
C 10	29.8	9.2	1589	6	US-10-750-185-28916
C 11	29.8	9.2	1912	6	US-10-750-185-52845
C 12	29.4	9.1	744	6	US-10-750-185-39160
C 13	29.4	9.1	3648	6	US-10-750-185-35726
C 14	28.8	8.9	1053	6	US-10-750-185-47371
C 15	28.6	8.9	1250	6	US-10-750-185-19975
C 16	28.6	8.9	1200	6	US-10-750-185-39178
C 17	28.6	8.9	1407	6	US-10-793-626-867
C 18	28.6	8.9	1407	6	US-10-793-626-1617
C 19	28.6	8.9	2563	6	US-10-750-185-35852
C 20	28.6	8.9	2580	6	US-10-793-626-4323
C 21	28.6	8.9	2856	6	US-10-793-626-3440
C 22	28.6	8.9	3227	6	US-10-793-626-3874
C 23	28.4	8.8	1605	6	US-10-750-185-53178

24	28.2	8.7	1436	6	US-10-750-185-60381	Sequence 60381, A
25	28.2	8.7	1706	6	US-10-750-185-38934	Sequence 38934, A
c 26	28.2	8.7	2066	6	US-10-750-185-57020	Sequence 57020, A
27	28.2	8.7	5143	6	US-10-750-185-55499	Sequence 55499, A
28	28	8.7	598	6	US-10-750-185-3967	Sequence 3967, Ap
29	28	8.7	854	6	US-10-750-185-42669	Sequence 42669, A
30	28	8.7	1370	6	US-10-750-185-25468	Sequence 25468, A
31	27.8	8.6	822	7	US-11-015-546A-36	Sequence 36, Appl
32	27.8	8.6	5286	6	US-10-750-185-45676	Sequence 45676, A
33	27.6	8.5	600	6	US-10-750-185-41520	Sequence 41520, A
34	27.6	8.5	740	6	US-10-750-185-55111	Sequence 55111, A
c 35	27.6	8.5	1144	6	US-10-750-185-29217	Sequence 29217, A
c 36	27.6	8.5	1191	6	US-10-750-185-45797	Sequence 45797, A
c 37	27.6	8.5	1264	6	US-10-750-185-28522	Sequence 28522, A
c 38	27.6	8.5	1442	6	US-10-750-185-25029	Sequence 25029, A
c 39	27.6	8.5	1499	6	US-10-750-185-57653	Sequence 57653, A
40	27.6	8.5	1768	6	US-10-750-185-61684	Sequence 61684, A
41	27.6	8.5	1861	6	US-10-750-185-36855	Sequence 36855, A
42	27.6	8.5	2044	6	US-10-750-185-39730	Sequence 39730, A
c 43	27.6	8.5	2306	6	US-10-750-185-37284	Sequence 37284, A
44	27.6	8.5	2825	6	US-10-750-185-56778	Sequence 56778, A
45	27.4	8.5	861	6	US-10-750-185-46047	Sequence 46047, A

ALIGNMENTS

RESULT 1  
US-11-112-908-45/c  
; Sequence 45, Application US/11112908  
; Publication No. US20050260659A1  
; GENERAL INFORMATION:  
; APPLICANT: Harris, Cole  
; APPLICANT: Davis, Lisa M.  
; TITLE OF INVENTION: Breast Cancer Biomarkers  
; FILE REFERENCE: 04-164-US  
; CURRENT APPLICATION NUMBER: US/11/112,908  
; CURRENT FILING DATE: 2005-04-22  
; PRIOR APPLICATION NUMBER: US 60/564,758  
; PRIOR FILING DATE: 2004-04-23  
; PRIOR APPLICATION NUMBER: US 60/575,978  
; PRIOR FILING DATE: 2004-06-01  
; PRIOR APPLICATION NUMBER: US 60/631,702  
; PRIOR FILING DATE: 2004-11-30  
; PRIOR APPLICATION NUMBER: US 60/633,826  
; PRIOR FILING DATE: 2004-12-07  
; NUMBER OF SEQ ID NOS: 511  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 45  
; LENGTH: 182314  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-11-112-908-45

Query Match	36.0%;	Score 116.2;	DB 7;	Length 182314;
Best Local Similarity	66.0%;	Pred. No. 5.2e-28;		
Matches	208;	Conservative	0;	Mismatches 88;
Indels	19;	Gaps	2;	
Qy	13	CTATGCTCCACGCCCAAGTGCCTGAACATCTGGGCGCTGATTTCTGGGCATCTCTCATGAC	72	
Db	27380	CCATGCTTCCACTGCAAAAGTGCCTGAACATCTATGCTCTGATCTTTCGGCATCTCTGTAGC	27321	
Qy	73	CATTGGATTACCTCTGTTACTGGTATTTCGGCTCTGTGACATCTACCATATATATGTTACA	132	
Db	27320	CACGTGCACTCATCATGATTTTCCACCACCTAGCTAGGAGATAATCTCCCAAGCAATTTCTCA	27261	
Qy	133	GATAATACAGAAAAACGGGCTTACTAGTAGCCGCCCATAGCTCTGCAACCTTTTGCACTCC	192	
Db	27260	GATGATTAAGATCTCCAGGCCAGTAGTTG-----TGCCCATATACTTCC	27216	
Qy	193	ACTGTGCAATGCTGGCCCTGCAC----GCTGGGGCTGTGCCCCCTTGGCTCTGCTG	248	
Db	27215	ACCGTCCATTGCTGGCCCTTGACCTTGCTGGGGCTGTGCCCCCTTGGCTCTGCTG	27156	



Matches 83; Conservative 2; Mismatches 18; Indels 1; Gaps 1;

Qy 2 GGGGCCAGGCTATGCTCCACC-GCCAACTGCTGAACATCTGGGCGCTGATCTGGG 60  
Db 174 GGGGCCAGGCTATGCTCCACCAGGCAAGTCTGACATGTGGGCGCTGATCTGGG 115  
Qy 61 CATCTCATGACCAATGGAATTCACCTGTTTACTGGTATTCGGCT 104  
Db 114 CATCTCATGACCAATTCGCTCATCGTCAATCCAGTGTGATCT 71

RESULT 5

US-10-750-185-59126  
; Sequence 59126, Application US/10750185  
; Publication No. US20050260603A1  
; GENERAL INFORMATION:  
; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MM1100-2  
; CURRENT APPLICATION NUMBER: US/10/750,185  
; CURRENT FILING DATE: 2003-12-31  
; PRIOR APPLICATION NUMBER: US 60/437,482  
; PRIOR FILING DATE: 2002-12-31  
; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: PatentIN version 3.1  
; SEQ ID NO 59126  
; LENGTH: 797  
; TYPE: DNA  
; ORGANISM: Bovine 19866880541216  
US-10-750-185-59126

Query Match 10.5%; Score 34; DB 6; Length 797;  
Best Local Similarity 59.2%; Pred. No. 0.092;  
Matches 58; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

Qy 184 TTGACTCCACTGTGCAATGCTGGCGCTGACGCTGGGGCTGTTCGCCCTGCCCTTGG 243  
Db 61 TTGAATCCACGACAGTCTTCACGAGGCTGAGGGCTGACCCCTTAAG 120  
Qy 244 TCTGCCCCATGATACAGAGTTTATACCAACACCT 281  
Db 121 GCTTCCTTCGTATCGGAAGCAATGCCCATCCACTT 158

RESULT 6

US-10-750-185-32680  
; Sequence 32680, Application US/10750185  
; Publication No. US20050260603A1  
; GENERAL INFORMATION:  
; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MM1100-2  
; CURRENT APPLICATION NUMBER: US/10/750,185  
; CURRENT FILING DATE: 2003-12-31  
; PRIOR APPLICATION NUMBER: US 60/437,482  
; PRIOR FILING DATE: 2002-12-31  
; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: PatentIN version 3.1  
; SEQ ID NO 32680  
; LENGTH: 3235  
; TYPE: DNA

; ORGANISM: Bovine 19866880710369  
US-10-750-185-32680

Query Match 10.1%; Score 32.6; DB 6; Length 3235;  
Best Local Similarity 51.7%; Pred. No. 0.54;  
Matches 74; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

Qy 48 CCTGATCTGGGCATCTCATGACCATTTGGATTCACCCCTGTACTGTATTCGGCTCTG 107  
Db 2853 CCTTCTCCAGGGGATCTTTGGACCTTAGGCATCAACCTGGGTCTCGCATGTTGCAGGC 2912  
Qy 108 TGACAGTCTACCATATTATCTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGC 167  
Db 2913 AGACACTTTTACCGTCTGAGCCACAGGAAAAAATGGTAACACATTATATATGGGCTGTGT 2972  
Qy 168 CCATAGCTCGCAACCTTTGCACT 190  
Db 2973 TCACATACAGAAAGATCTGATCT 2995

RESULT 7

US-10-750-185-51883/c  
; Sequence 51883, Application US/10750185  
; Publication No. US20050260603A1  
; GENERAL INFORMATION:  
; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MM1100-2  
; CURRENT APPLICATION NUMBER: US/10/750,185  
; CURRENT FILING DATE: 2003-12-31  
; PRIOR FILING DATE: 2002-12-31  
; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: PatentIN version 3.1  
; SEQ ID NO 51883  
; LENGTH: 3085  
; TYPE: DNA  
; ORGANISM: Bovine 19866880919936  
US-10-750-185-51883

Query Match 9.6%; Score 31; DB 6; Length 3085;  
Best Local Similarity 55.0%; Pred. No. 1.8;  
Matches 61; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

Qy 49 CCTGATCTGGGCATCTCATGACCATTTGGATTCACCCCTGTACTGTGTTATTCGGCTCTGT 108  
Db 2027 CCTTCTCAGAGGATCTTCTGACCCAGGGATTTGAACTGGGACTCTCTGTTTGCAGGCA 1968  
Qy 109 GACAGTCTACCATATTATTTACAGATAATACAGAAAAACGGGGTTACTA 159  
Db 1967 GATTCCTTTACCACTTAGCTACCAGGGAAGCCCTGTTAGGTCTTAACAA 1917

RESULT 8

US-10-750-185-55588/c  
; Sequence 55588, Application US/10750185  
; Publication No. US20050260603A1  
; GENERAL INFORMATION:  
; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MM1100-2

; CURRENT APPLICATION NUMBER: US/10/750,185  
; CURRENT FILING DATE: 2003-12-31  
; PRIOR APPLICATION NUMBER: US 60/437,482  
; PRIOR FILING DATE: 2002-12-31  
; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: PatentIN version 3.1  
; SEQ ID NO 55588  
; LENGTH: 1648  
; TYPE: DNA  
; ORGANISM: Bovine 1986688068347  
US-10-750-185-55588

Query Match 9.4%; Score 30.4; DB 6; Length 1648;  
Best Local Similarity 59.1%; Pred. No. 2.1; Mismatches 0; Indels 0; Gaps 0;  
Matches 52; Conservative 0

Qy 105 CTGTGACAGTCTACCATATTATGTTACAGATAATACAGGAAACGGGGTTACTAGTAGC 164  
Db 1406 CAGAGAGAGATGAATCTCTATAAACTACAAATAAGGAAAGAGCTGTTAGAAAAAGC 1347

Qy 165 CGCCATAGCTGCAACCTTGCCTCC 192  
Db 1346 AGTCTCGCTGCAGCCTTTTCACCCC 1319

## RESULT 9

US-10-750-185-30919  
; Sequence 30919, Application US/10750185  
; Publication No. US20050260603A1  
; GENERAL INFORMATION:  
; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MM1100-2  
; CURRENT APPLICATION NUMBER: US/10/750,185  
; CURRENT FILING DATE: 2003-12-31  
; PRIOR APPLICATION NUMBER: US 60/437,482  
; PRIOR FILING DATE: 2002-12-31  
; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: PatentIN version 3.1  
; SEQ ID NO 30919  
; LENGTH: 1242  
; TYPE: DNA  
; ORGANISM: Bovine 19866881328575  
US-10-750-185-30919

Query Match 9.3%; Score 30; DB 6; Length 1242;  
Best Local Similarity 48.8%; Pred. No. 2.5; Mismatches 0; Indels 0; Gaps 0;  
Matches 81; Conservative 0

Qy 42 TCTGGCCCTGATCTGGGCACTCTCATGACCATGATTCACCTGTTACTGTTATTCG 101  
Db 1074 TCTGGAGCCTTGGCTGGGTTTCACTGGGAGAAATTTGATGATGACTGGTCTACAAGATGG 1133

Qy 102 GCTCTGTGACAGTCTACCATATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGT 161  
Db 1134 GATCTGTATCTGTGAAGTGTGTGGCTCTTGACATACAGAGGATGTGGGATGCTCT 1193

Qy 162 AGCCGCCCATAGCTGCACCTTTGCACTCCACTGCAATGCTGG 207  
Db 1194 GGCTTCCCTCTTCTCTTGCCTCTTCACTGCTCCCTCAGAACCTCTCG 1239

## RESULT 10

US-10-750-185-28916/c  
; Sequence 28916, Application US/10750185  
; Publication No. US20050260603A1  
; GENERAL INFORMATION:

; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MM1100-2  
; CURRENT APPLICATION NUMBER: US/10/750,185  
; CURRENT FILING DATE: 2003-12-31  
; PRIOR APPLICATION NUMBER: US 60/437,482  
; PRIOR FILING DATE: 2002-12-31  
; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: PatentIN version 3.1  
; SEQ ID NO 28916  
; LENGTH: 1589  
; TYPE: DNA  
; ORGANISM: Bovine 19866880854078  
US-10-750-185-28916

Query Match 9.2%; Score 29.8; DB 6; Length 1589;  
Best Local Similarity 55.2%; Pred. No. 3.3; Mismatches 0; Indels 0; Gaps 0;  
Matches 58; Conservative 0

Qy 156 ACTAGTAGCCGCCCATAGCTGCAACCTTTGCACTCCACTGTGCAATGCTGGCCCTGCAC 215  
Db 1588 ACTTTTAGGAGCCTGAATCCAGGAGAGGAAAGCCCTCAGTGTGTCATGACCTGCC 1529

Qy 216 GCTGGGGCTGTTGCCCTGCCCTGGTCTGCTGCCCTAGATACA 260  
Db 1528 CGTATGGCTGCTTCTCCCGAGACCTGGTTCGTCCTCCTCACCCTTCA 1484

## RESULT 11

US-10-750-185-52845  
; Sequence 52845, Application US/10750185  
; Publication No. US20050260603A1  
; GENERAL INFORMATION:  
; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MM1100-2  
; CURRENT APPLICATION NUMBER: US/10/750,185  
; CURRENT FILING DATE: 2003-12-31  
; PRIOR APPLICATION NUMBER: US 60/437,482  
; PRIOR FILING DATE: 2002-12-31  
; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: PatentIN version 3.1  
; SEQ ID NO 52845  
; LENGTH: 1912  
; TYPE: DNA  
; ORGANISM: Bovine 19866880622237  
US-10-750-185-52845

Query Match 9.2%; Score 29.8; DB 6; Length 1912;  
Best Local Similarity 52.9%; Pred. No. 3.6; Mismatches 0; Indels 0; Gaps 0;  
Matches 64; Conservative 0

Qy 181 CCTTTGCACTCCCACTGTGCAATGCTGGCCCTGCACGCTGGGGCTGTGGCCCTGCCCT 240  
Db 1572 CCTTCTCCTCCTTGAGAATCTTGACCTTATCTCTCTTGAGAGCCCGACACCCCT 1631

Qy 241 TGGTCTGCCCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCTTCAAT 300  
Db 1632 TTCTCTCCTCCTCAGTGACCCCTGGACTTTTCTCAAGCTGCTCCTAGGAGTGACTCTCTCAAT 1691

Qy 301 A 301

```
Db      1692 A 1692

RESULT 12
US-10-750-185-39160
; Sequence 39160, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 39160
; LENGTH: 744
; TYPE: DNA
; ORGANISM: Bovine 19866881456696
US-10-750-185-39160

Query Match      9.1%; Score 29.4; DB 6; Length 744;
Best Local Similarity 56.8%; Pred. No. 3;
Matches 54; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

Qy      4 GGCCAGGCTATGCTCCTCACCGCAAGTGCCTGAACATCTGGGCCCTGATCTGGGGCAT 63
Db      30 GTCTGGGACTGTGTCAGCCCAAGATATCTGAATGACTGCCATGCCCTCTCCAGGGCAT 89

Qy      54 CCTCATGACCATGATTCACCCCTGTTACTGGTAT 98
Db      90 CTTCCTGACCCAGAGATCAAAACCCGTCCTCTAT 124

RESULT 13
US-10-750-185-35726
; Sequence 35726, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 35726
; LENGTH: 3648
; TYPE: DNA
; ORGANISM: Bovine 19866880517544
US-10-750-185-35726

Query Match      9.1%; Score 29.4; DB 6; Length 3648;
Best Local Similarity 52.9%; Pred. No. 6.7;
Matches 63; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

Db      1692 A 1692

Qy      42 TCTGGGCCCTGATTTCTGGGCATCTCTATGACCATTTGGATTCACCCCTGTTACTGTGATTCG 101
Db      2749 TCTATGCCCTCTCTCCAGGGGATCTTCTGACCCAGGGATTTGAACCCAGGTCTCCACATT 2808

Qy      102 GCTCTGTGACAGTCTACCATATTTATGTTACAGATAATACAGGAAAAACGGGGTTACTAG 160
Db      2809 GCAGGTGGATTTCTTTACCATCTGAGCCACCAGGGAAACCCAAATACTGGAGTGGGTAG 2867

RESULT 14
US-10-750-185-47371
; Sequence 47371, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 47371
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Bovine 19866880909469
US-10-750-185-47371

Query Match      8.9%; Score 28.8; DB 6; Length 1053;
Best Local Similarity 58.0%; Pred. No. 5.7;
Matches 51; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

Qy      10 GGCCATGCTCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCGGGCATCCCTCAT 69
Db      936 GGCTTCTCCAGGCAATAATCTGGAGTGGATCACCAGGCCCTGCTTCAGGGGATCTTCCT 995

Qy      70 GACCATTTGGATTCACCCCTGTTACTGGTA 97
Db      996 GACTCAGGATTTGAACCTGCACTCCTA 1023

RESULT 15
US-10-750-185-19975
; Sequence 19975, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19975
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Bovine MMBT04626
US-10-750-185-19975
```

Query Match 8.9%; Score 28.6; DB 6; Length 600;  
Best Local Similarity 64.2%; Pred. No. 5.1; 24; Indels 0; Gaps 0;  
Matches 43; Conservative 0; Mismatches 24; Indels 0; Gaps 0;  
Qy 103 CTCTGTGACAGTCTTACCATATTTATGTTACAGATAATACAGGAAAAACGGGTTACTAGTA 162  
Db 150 CTGTGGGACAGTCTTCGCTTCCACGAAACAGAGAGAGGAGGAAAAATTCATTACTGGAT 209  
Qy 163 GCCGCCC 169  
Db 210 GCTGCCC 216

Search completed: December 6, 2005, 23:16:52  
Job time : 221.487 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 21:43:57 ; Search time 3686.8 Seconds  
(without alignments)  
4099.008 Million cell updates/sec

Title: US-09-980-046B-6  
Perfect score: 323  
Sequence: 1 cggggccagcctatgcct.....gtgcagctgttgaaaaa 323

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 41078325 seqs, 23393541228 residues  
Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_hcc:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_est7:\*  
9: gb\_ges1:\*  
10: gb\_ges2:\*  
11: gb\_ges3:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	321.4	99.5	365	2	BG194579 RST13741
2	321.4	99.5	365	2	BG218402 RST18270
3	321.4	99.5	366	2	BG218401 RST38269
4	321.4	99.5	367	2	BG186275 RST5236 A
5	321.4	99.5	367	2	BG188886 RST7920 A
6	321.4	99.5	367	2	BG203761 RST23136
7	321.4	99.5	370	2	BG192020 RST11127
8	321.4	99.5	414	3	BM662200 UI-B-C10-
9	321.4	99.5	469	7	CN264982 170005999
10	321.4	99.5	498	1	AW574932 UI-HF-BL0
11	321.4	99.5	498	6	C3420737 UI-H-FGO-
12	321.4	99.5	504	5	BQ447972 UI-H-EU1-
13	321.4	99.5	505	3	BM672499 UI-B-CQ1-
14	321.4	99.5	510	3	BM767351 K-EST0049
15	321.4	99.5	522	5	BU783190 in04d02.x
16	321.4	99.5	552	6	CA414261 UI-H-E20-
17	321.4	99.5	563	7	CT001900 CT001900
18	321.4	99.5	567	5	BQ776064 UI-H-FR0-
19	321.4	99.5	568	5	BQ706056 AGENCOURT
20	321.4	99.5	574	5	BU731878 UI-B-C11-
21	321.4	99.5	587	5	BU680316 UI-CF-DU1
22	321.4	99.5	588	5	BU731736 UI-B-C11-

C 23	321.4	99.5	589	6	CA413789
C 24	321.4	99.5	611	6	CA428144
C 25	321.4	99.5	614	5	BQ440412
C 26	321.4	99.5	629	7	CR983622
C 27	321.4	99.5	635	7	CR985541
C 28	321.4	99.5	641	7	CR982834
C 29	321.4	99.5	645	6	CA414705
C 30	321.4	99.5	649	3	BQ181165
C 31	321.4	99.5	651	8	DR155863
C 32	321.4	99.5	652	8	DR158870
C 33	321.4	99.5	654	3	BQ181609
C 34	321.4	99.5	655	8	DR159276
C 35	321.4	99.5	656	5	BU686429
C 36	321.4	99.5	670	3	BM995109
C 37	321.4	99.5	674	7	CN264991
C 38	321.4	99.5	675	8	DR158734
C 39	321.4	99.5	677	6	CA421204
C 40	321.4	99.5	683	6	CA419982
C 41	321.4	99.5	683	8	CX165582
C 42	321.4	99.5	691	3	BQ219055
C 43	321.4	99.5	704	6	CD173064
C 44	321.4	99.5	706	6	CA429526
C 45	321.4	99.5	710	6	CD172976

ALIGNMENTS

RESULT 1  
BG194579  
LOCUS BG194579 365 bp mRNA linear EST 21-APR-2001  
DEFINITION RST13741 Athersys RAGE Library Homo sapiens cdna, mRNA sequence.  
ACCESSION BG194579  
VERSION BG194579.1 GI:13716266  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
REFERENCE 1 (bases 1 to 365)  
AUTHORS Harrington,J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,  
Cain,S., Leventhal,C., Thornton,M., Ramachandran,R.,  
Whittington,J., Lerner,L., Costanzo,D., McElligott,K., Booser,S.,  
Mays,R., Smith,R., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K.,  
Offenbacher,J., Danzig,J. and Ducar,M.  
TITLE Creation of genome-wide protein expression libraries using random  
activation of gene expression  
JOURNAL Nat. Biotechnol. 19 (5), 440-445 (2001)  
PUBMED 11329013  
COMMENT Contact: Scott J. Cain  
Athersys, Inc.  
3201 Carnegie Ave, Cleveland, OH 44115, USA  
Tel: 216 431 9900  
Fax: 216 361 9596  
Email: scain@atersys.com  
High quality sequence stop: 365.  
Location/Qualifiers  
1. .365  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/cell\_line="HT1080"  
/clone\_lib="Athersys RAGE Library"  
/note="See 'Creation of Genome-wide Protein Expression  
Libraries using Random Activation of Gene Expression',  
Nature Biotechnology, in press. Note that even though the  
cell type indicated is HT1080, since a random activation  
method was used, these sequence tags are not necessarily  
expressed in HT1080 under normal circumstances."

ORIGIN

Query Match 99.5%; Score 321.4; DB 2; Length 365;

Best Local Similarity 99.7%; Pred. No. 1.3e-81;  
Matches 322; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGGGCCAGCCCTATGCCCTCCACCGCAAGTGCTGAAACATCTGGGCCCTGATTCCTGGG 60  
Db 35 CGGGGCCAGCCCTATGCCCTCCACCGCAAGTGCTGAAACATCTGGGCCCTGATTCCTGGG 94

Qy 61 CATCCTCATGACCAATTCGATTTCACCCCTGTTACTGGTATTTCGGCTCTGTGACAGTCTACCA 120  
Db 95 CATCCTCATGACCAATTCGATTTCACCCCTGTTACTGGTATTTCGGCTCTGTGACAGTCTACCA 154

Qy 121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCCTGCAA 180  
Db 155 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCCTGCAA 214

Qy 181 CTTTGGCACTCCACTGTGCAATGCTGGCCCTGCAAGTGGGGCTGTTGCCCTGCCCCCT 240  
Db 215 CTTTGGCACTCCACTGTGCAATGCTGGCCCTGCAAGTGGGGCTGTTGCCCTGCCCCCT 274

Qy 241 TGGTCCCTGCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 300  
Db 275 TGGTCCCTGCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 334

Qy 301 AAAGTGCACGTGCTTGTGAAAAA 323  
Db 335 AAAGTGCACGTGCTTGTGAAAAA 357

RESULT 2  
BG218402  
LOCUS BG218402 365 bp mRNA linear EST 21-APR-2001  
DEFINITION RST38270 Atherys RAGE Library Homo sapiens cDNA, mRNA sequence.  
ACCESSION BG218402  
VERSION BG218402.1 GI:13744551  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE 1 (bases 1 to 365)  
Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,  
Cain,S., Leventhal,C., Thornton,M., Ramachandran,R.,  
Whittington,J., Lerner,L., Costanzo,D., McElligott,K., Booser,S.,  
Mays,R., Smith,E., Veloso,N., Klika,A., Hess,J., Cochren,K., Lo,K.,  
Offenbacher,J., Danzig,J. and Ducar,M.  
Creation of genome-wide protein expression libraries using random  
activation of gene expression  
Nat. Biotechnol. 19 (5), 440-445 (2001)  
JOURNAL PUBMED 11329013  
COMMENT Contact: Scott J. Cain  
Atherys, Inc.  
3201 Carnegie Ave, Cleveland, OH 44115, USA  
Tel: 216 431 9900  
Fax: 216 361 9596  
Email: scain@atherys.com  
High quality sequence stop: 365.  
Location/Qualifiers  
FEATURES  
source  
1. .365  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/cell\_line="HT1080"  
/clone\_lib="Atherys RAGE Library"  
/note="See 'Creation of Genome-wide Protein Expression  
Libraries using Random Activation of Gene Expression',  
Nature Biotechnology, in press. Note that even though the  
cell type indicated is HT1080, since a random activation  
method was used, these sequence tags are not necessarily  
expressed in HT1080 under normal circumstances."

Query Match 99.5%; Score 321.4; DB 2; Length 365;

ORIGIN

Best Local Similarity 99.7%; Pred. No. 1.3e-81;  
Matches 322; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGGGCCAGCCCTATGCCCTCCACCGCAAGTGCTGAAACATCTGGGCCCTGATTCCTGGG 60  
Db 35 CGGGGCCAGCCCTATGCCCTCCACCGCAAGTGCTGAAACATCTGGGCCCTGATTCCTGGG 94

Qy 61 CATCCTCATGACCAATTCGATTTCACCCCTGTTACTGGTATTTCGGCTCTGTGACAGTCTACCA 120  
Db 95 CATCCTCATGACCAATTCGATTTCACCCCTGTTACTGGTATTTCGGCTCTGTGACAGTCTACCA 154

Qy 121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCCTGCAA 180  
Db 155 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCCTGCAA 214

Qy 181 CTTTGGCACTCCACTGTGCAATGCTGGCCCTGCAAGTGGGGCTGTTGCCCTGCCCCCT 240  
Db 215 CTTTGGCACTCCACTGTGCAATGCTGGCCCTGCAAGTGGGGCTGTTGCCCTGCCCCCT 274

Qy 241 TGGTCCCTGCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 300  
Db 275 TGGTCCCTGCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 334

Qy 301 AAAGTGCACGTGCTTGTGAAAAA 323  
Db 335 AAAGTGCACGTGCTTGTGAAAAA 357

RESULT 3  
BG218401  
LOCUS BG218401 366 bp mRNA linear EST 21-APR-2001  
DEFINITION RST38269 Atherys RAGE Library Homo sapiens cDNA, mRNA sequence.  
ACCESSION BG218401  
VERSION BG218401.1 GI:13744550  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE 1 (bases 1 to 366)  
Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,  
Cain,S., Leventhal,C., Thornton,M., Ramachandran,R.,  
Whittington,J., Lerner,L., Costanzo,D., McElligott,K., Booser,S.,  
Mays,R., Smith,E., Veloso,N., Klika,A., Hess,J., Cochren,K., Lo,K.,  
Offenbacher,J., Danzig,J. and Ducar,M.  
Creation of genome-wide protein expression libraries using random  
activation of gene expression  
Nat. Biotechnol. 19 (5), 440-445 (2001)  
JOURNAL PUBMED 11329013  
COMMENT Contact: Scott J. Cain  
Atherys, Inc.  
3201 Carnegie Ave, Cleveland, OH 44115, USA  
Tel: 216 431 9900  
Fax: 216 361 9596  
Email: scain@atherys.com  
High quality sequence stop: 366.  
Location/Qualifiers  
FEATURES  
source  
1. .366  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/cell\_line="HT1080"  
/clone\_lib="Atherys RAGE Library"  
/note="See 'Creation of Genome-wide Protein Expression  
Libraries using Random Activation of Gene Expression',  
Nature Biotechnology, in press. Note that even though the  
cell type indicated is HT1080, since a random activation  
method was used, these sequence tags are not necessarily  
expressed in HT1080 under normal circumstances."

Query Match 99.5%; Score 321.4; DB 2; Length 366;

ORIGIN

Best Local Similarity 99.7%; Pred. No. 1.3e-81;  
Matches 322; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGGGCCAGGCGCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCGGG 60  
Db 36 CGGGGCCAGGCGCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCGGG 95

Qy 61 CATCCTCATGACCAATGGATTCAACCCCTGTTACTGGTATTCGGCTCTGTGACAGTCTACCA 120  
Db 96 CATCCTCATGACCAATGGATTCAACCCCTGTTACTGGTATTCGGCTCTGTGACAGTCTACCA 155

Qy 121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCCTGCAA 180  
Db 156 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCCTGCAA 215

Qy 181 CTTTGGCACTCCACTGTGCAATGCTGGCCCTGCGACGCTGGGGCTGTGGCCCTGCCCT 240  
Db 216 CTTTGGCACTCCACTGTGCAATGCTGGCCCTGCGACGCTGGGGCTGTGGCCCTGCCCT 275

Qy 241 TGGTCTGCCCTAGATACAGCAGTTTATATACCCACACACCTGTCTACAGTGTCAATCAAT 300  
Db 276 TGGTCTGCCCTAGATACAGCAGTTTATATACCCACACACCTGTCTACAGTGTCAATCAAT 335

Qy 301 AAAGTGCACGTGCTTGTGAAAAA 323  
Db 336 AAAGTGCACGTGCTTGTGAAAAA 358

RESULT 4  
BG186275  
LOCUS  
DEFINITION  
RST5236 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.  
ACCESSION  
BG186275  
VERSION  
BG186275.1 GI:13707962  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
REFERENCE  
1 (bases 1 to 367)  
Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,  
Cain,S., Leventhal,C., Thornton,M., Ramachandran,R.,  
Whittington,J., Lerner,L., Costanzo,D., McElligott,K., Booser,S.,  
Mays,R., Smith,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K.,  
Offenbacher,J., Danzig,J. and Ducar,M.  
Creation of genome-wide protein expression libraries using random  
activation of gene expression  
Nat. Biotechnol. 19 (5), 440-445 (2001)  
11329013  
Contact: Scott J. Cain  
Athersys, Inc.  
3201 Carnegie Ave, Cleveland, OH 44115, USA  
Tel: 216 431 9900  
Fax: 216 361 9596  
Email: scain@athersys.com  
High quality sequence stop: 367.  
Location/Qualifiers  
1..367  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/cell\_line="HT1080"  
/clone\_lib="Athersys RAGE Library"  
/note="See 'Creation of Genome-wide Protein Expression  
Libraries using Random Activation of Gene Expression',  
Nature Biotechnology, in press. Note that even though the  
cell type indicated is HT1080, since a random activation  
method was used, these sequence tags are not necessarily  
expressed in HT1080 under normal circumstances."

ORIGIN

Query Match 99.5%; Score 321.4; DB 2; Length 367;

Best Local Similarity 99.7%; Pred. No. 1.3e-81;  
Matches 322; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGGGCCAGGCGCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCGGG 60  
Db 37 CGGGGCCAGGCGCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCGGG 96

Qy 61 CATCCTCATGACCAATGGATTCAACCCCTGTTACTGGTATTCGGCTCTGTGACAGTCTACCA 120  
Db 97 CATCCTCATGACCAATGGATTCAACCCCTGTTACTGGTATTCGGCTCTGTGACAGTCTACCA 156

Qy 121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCCTGCAA 180  
Db 157 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCCTGCAA 216

Qy 181 CTTTGGCACTCCACTGTGCAATGCTGGCCCTGCGACGCTGGGGCTGTGGCCCTGCCCT 240  
Db 217 CTTTGGCACTCCACTGTGCAATGCTGGCCCTGCGACGCTGGGGCTGTGGCCCTGCCCT 276

Qy 241 TGGTCTGCCCTAGATACAGCAGTTTATATACCCACACACCTGTCTACAGTGTCAATCAAT 300  
Db 277 TGGTCTGCCCTAGATACAGCAGTTTATATACCCACACACCTGTCTACAGTGTCAATCAAT 336

Qy 301 AAAGTGCACGTGCTTGTGAAAAA 323  
Db 337 AAAGTGCACGTGCTTGTGAAAAA 359

RESULT 5  
BG188886  
LOCUS  
DEFINITION  
RST7920 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.  
ACCESSION  
BG188886  
VERSION  
BG188886.1 GI:13710573  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
REFERENCE  
1 (bases 1 to 367)  
Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,  
Cain,S., Leventhal,C., Thornton,M., Ramachandran,R.,  
Whittington,J., Lerner,L., Costanzo,D., McElligott,K., Booser,S.,  
Mays,R., Smith,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K.,  
Offenbacher,J., Danzig,J. and Ducar,M.  
Creation of genome-wide protein expression libraries using random  
activation of gene expression  
Nat. Biotechnol. 19 (5), 440-445 (2001)  
11329013  
Contact: Scott J. Cain  
Athersys, Inc.  
3201 Carnegie Ave, Cleveland, OH 44115, USA  
Tel: 216 431 9900  
Fax: 216 361 9596  
Email: scain@athersys.com  
High quality sequence stop: 367.  
Location/Qualifiers  
1..367  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/cell\_line="HT1080"  
/clone\_lib="Athersys RAGE Library"  
/note="See 'Creation of Genome-wide Protein Expression  
Libraries using Random Activation of Gene Expression',  
Nature Biotechnology, in press. Note that even though the  
cell type indicated is HT1080, since a random activation  
method was used, these sequence tags are not necessarily  
expressed in HT1080 under normal circumstances."

ORIGIN

Query Match 99.5%; Score 321.4; DB 2; Length 367;

Best Local Similarity 99.7%; Pred. No. 1.3e-81;  
Matches 322; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGGGCCAGCCCTATGCTCCACCGCAAGTGCTGAAACATCTGGGCCCTGATTCGGG 60  
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
37 CGGGGCCAGCCCTATGCTCCACCGCAAGTGCTGAAACATCTGGGCCCTGATTCGGG 96  
Qy 61 CATCCTCATGACCAATTCGATTCACCCCTGTTACTGGTATTCGGCTCTGTGACAGTCTACCA 120  
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
97 CATCCTCATGACCAATTCGATTCACCCCTGTTACTGGTATTCGGCTCTGTGACAGTCTACCA 156  
Qy 121 TATTATGTTACAGATAATACAGGAAAAACCGGGTTACTAGTAGCCGCCCATAGCCTGCAA 180  
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
157 TATTATGTTACAGATATACAGGAAAAACCGGGTTACTAGTAGCCGCCCATAGCCTGCAA 216  
Qy 181 CTTTGTCACTCCACTGTGCAATGTGCGCCCTGCAAGTGGGGCTGTGGCCCTGCCCCCT 240  
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
217 CTTTGTCACTCCACTGTGCAATGTGCGCCCTGCAAGTGGGGCTGTGGCCCTGCCCCCT 276  
Qy 241 TGGTCTGCCCCCTAGATACAGCAGTTTATATCCACACACCTGTCTACAGTGTCAATCAAT 300  
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
277 TGGTCTGCCCCCTAGATACAGCAGTTTATATCCACACACCTGTCTACAGTGTCAATCAAT 336  
Qy 301 AAAGTGCACGTGCTTGTGAAAAA 323  
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
337 AAAGTGCACGTGCTTGTGAAAAA 359

RESULT 6  
BG192020  
LOCUS BG192020 367 bp mRNA linear EST 21-APR-2001  
DEFINITION RST23136 Atherys RAGE Library Homo sapiens cDNA, mRNA sequence.  
ACCESSION BG203761  
VERSION BG203761.1 GI:13725432  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE 1 (bases 1 to 367)  
AUTHORS Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,  
Cain,S., Leventhal,C., Thornton,M., Ramachandran,R.,  
Whittington,J., Lerner,L., Costanzo,D., McElligott,K., Booser,S.,  
Mays,R., Smith,E., Veloso,N., Klika,A., Hess,J., Cochren,K., Lo,K.,  
Offenbacher,J., Danzig,J. and Ducar,M.  
TITLE Creation of genome-wide protein expression libraries using random  
activation of gene expression  
JOURNAL Nat. Biotechnol. 19 (5), 440-445 (2001)  
PUBMED 11329013  
COMMENT Contact: Scott J. Cain  
Athersys, Inc.  
3201 Carnegie Ave, Cleveland, OH 44115, USA  
Tel: 216 431 9900  
Fax: 216 361 9596  
Email: scain@atersys.com  
High quality sequence stop: 367.  
Location/Qualifiers  
FEATURES  
Source  
1. .367  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone\_lib="HT1080"  
/notes="See 'Creation of Genome-wide Protein Expression  
Libraries using Random Activation of Gene Expression',  
Nature Biotechnology, in press. Note that even though the  
cell type indicated is HT1080, since a random activation  
method was used, these sequence tags are not necessarily  
expressed in HT1080 under normal circumstances."

Query Match 99.5%; Score 321.4; DB 2; Length 367;

ORIGIN

Best Local Similarity 99.7%; Pred. No. 1.3e-81;  
Matches 322; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGGGCCAGCCCTATGCTCCACCGCAAGTGCTGAAACATCTGGGCCCTGATTCGGG 60  
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
37 CGGGGCCAGCCCTATGCTCCACCGCAAGTGCTGAAACATCTGGGCCCTGATTCGGG 96  
Qy 61 CATCCTCATGACCAATTCGATTCACCCCTGTTACTGGTATTCGGCTCTGTGACAGTCTACCA 120  
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
97 CATCCTCATGACCAATTCGATTCACCCCTGTTACTGGTATTCGGCTCTGTGACAGTCTACCA 156  
Qy 121 TATTATGTTACAGATAATACAGGAAAAACCGGGTTACTAGTAGCCGCCCATAGCCTGCAA 180  
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
157 TATTATGTTACAGATATACAGGAAAAACCGGGTTACTAGTAGCCGCCCATAGCCTGCAA 216  
Qy 181 CTTTGTCACTCCACTGTGCAATGTGCGCCCTGCAAGTGGGGCTGTGGCCCTGCCCCCT 240  
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
217 CTTTGTCACTCCACTGTGCAATGTGCGCCCTGCAAGTGGGGCTGTGGCCCTGCCCCCT 276  
Qy 241 TGGTCTGCCCCCTAGATACAGCAGTTTATATCCACACACCTGTCTACAGTGTCAATCAAT 300  
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
277 TGGTCTGCCCCCTAGATACAGCAGTTTATATCCACACACCTGTCTACAGTGTCAATCAAT 336  
Qy 301 AAAGTGCACGTGCTTGTGAAAAA 323  
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
337 AAAGTGCACGTGCTTGTGAAAAA 359

RESULT 7  
BG192020  
LOCUS BG192020 370 bp mRNA linear EST 21-APR-2001  
DEFINITION RST11127 Atherys RAGE Library Homo sapiens cDNA, mRNA sequence.  
ACCESSION BG192020  
VERSION BG192020.1 GI:13713707  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE 1 (bases 1 to 370)  
AUTHORS Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,  
Cain,S., Leventhal,C., Thornton,M., Ramachandran,R.,  
Whittington,J., Lerner,L., Costanzo,D., McElligott,K., Booser,S.,  
Mays,R., Smith,E., Veloso,N., Klika,A., Hess,J., Cochren,K., Lo,K.,  
Offenbacher,J., Danzig,J. and Ducar,M.  
TITLE Creation of genome-wide protein expression libraries using random  
activation of gene expression  
JOURNAL Nat. Biotechnol. 19 (5), 440-445 (2001)  
PUBMED 11329013  
COMMENT Contact: Scott J. Cain  
Athersys, Inc.  
3201 Carnegie Ave, Cleveland, OH 44115, USA  
Tel: 216 431 9900  
Fax: 216 361 9596  
Email: scain@atersys.com  
High quality sequence stop: 370.  
Location/Qualifiers  
FEATURES  
Source  
1. .370  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone\_lib="HT1080"  
/notes="See 'Creation of Genome-wide Protein Expression  
Libraries using Random Activation of Gene Expression',  
Nature Biotechnology, in press. Note that even though the  
cell type indicated is HT1080, since a random activation  
method was used, these sequence tags are not necessarily  
expressed in HT1080 under normal circumstances."

Query Match 99.5%; Score 321.4; DB 2; Length 370;

ORIGIN

Best Local Similarity 99.7%; Pred. No. 1.3e-81; Mismatches 0; Gaps 0; Matches 322; Conservative 0; Indels 1; Gaps 0;

Qy 1 CGGGGCCAGCCCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCGGG 60  
Db 40 CGGGGCCAGCCCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCGGG 99

Qy 61 CATCCTCATGACCATTTGGATTTCACCTGTTACTGGTATTTCGGCTCTGTGACAGTCTACCA 120  
Db 100 CATCCTCATGACCATTTGGATTTCACCTGTTACTGGTATTTCGGCTCTGTGACAGTCTACCA 159

Qy 121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCTGCAG 180  
Db 160 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCTGCAG 219

Qy 181 CCTTTGCACTCCACTGTCATGCTGGCCCGACGCTGGGGCTGTGGCCCTGCCCT 240  
Db 220 CCTTTGCACTCCACTGTCATGCTGGCCCGACGCTGGGGCTGTGGCCCTGCCCT 279

Qy 241 TGGTCTGCCCTTAGATACAGCAGTTTATATACCCACACACCTGTCTACAGTGTCAATCAAT 300  
Db 280 TGGTCTGCCCTTAGATACAGCAGTTTATATACCCACACACCTGTCTACAGTGTCAATCAAT 339

Qy 301 AAAGTGCACTGCTGTGTGAAAAA 323  
Db 340 AAAGTGCACTGCTGTGTGAAAAA 362

RESULT 8  
BM662200/c  
LOCUS  
DEFINITION  
UI-E-CL0-acd-g-01-0-UI-s1 UI-E-CL0 Homo sapiens cDNA clone  
UI-E-CL0-acd-g-01-0-UI 3', mRNA sequence.  
ACCSSION  
BM662200  
VERSION  
BM662200.1 GI:18966548  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 414)  
Bonaldo,M.F., Lennon,G. and Soares,M.B.  
Normalization and subtraction: two approaches to facilitate gene  
discovery  
Genome Res. 6 (9), 791-806 (1996)  
889548  
JOURNAL  
PUBMED  
COMMENT  
Contact: Soares, MB  
Coordinated Laboratory for Computational Genomics  
University of Iowa  
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: bento-soares@uiowa.edu  
Tissue Procurement: Dr. Gregg Hageman  
CNA Library preparation: Dr. M. Bento Soares, University of Iowa  
CNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Researchers may obtain clones from Research  
Genetics (www.resgen.com).  
Seq primer: M13 Forward  
POLYA=Yes.  
Location/Qualifiers  
1..414  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="UI-E-CL0-acd-g-01-0-UI"  
/tissue\_type="human retina"  
/dev\_stage="adult"  
/lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
/clone\_lib="UI-E-CL0"  
/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a

modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-E-CL0 is a cDNA library containing the following tissue(s): retina. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The first-strand cDNA contains a library tag sequence of oligonucleotide used to prime the synthesis of located between the Not I site and the (dT)18 tail. The sequence tag for this library is CCGCG. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI). TAG TISSUE=human retina TAG\_L18=UI-E-CL0 TAG\_SEQ=CCGCG"

ORIGIN

Query Match 99.5%; Score 321.4; DB 3; Length 414;  
Best Local Similarity 99.7%; Pred. No. 1.3e-81;  
Matches 322; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGGGCCAGCCCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCGGG 60  
Db 336 CGGGGCCAGCCCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCGGG 277

Qy 61 CATCCTCATGACCATTTGGATTTCACCTGTTACTGGTATTTCGGCTCTGTGACAGTCTACCA 120  
Db 276 CATCCTCATGACCATTTGGATTTCACCTGTTACTGGTATTTCGGCTCTGTGACAGTCTACCA 217

Qy 121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCTGCAG 180  
Db 216 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCTGCAG 157

Qy 181 CCTTTGCACTCCACTGTCATGCTGGCCCGACGCTGGGGCTGTGGCCCTGCCCT 240  
Db 156 CCTTTGCACTCCACTGTCATGCTGGCCCGACGCTGGGGCTGTGGCCCTGCCCT 97

Qy 241 TGGTCTGCCCTTAGATACAGCAGTTTATATACCCACACACCTGTCTACAGTGTCAATCAAT 300  
Db 96 TGGTCTGCCCTTAGATACAGCAGTTTATATACCCACACACCTGTCTACAGTGTCAATCAAT 37

Qy 301 AAAGTGCACTGCTGTGTGAAAAA 323  
Db 36 AAAGTGCACTGCTGTGTGAAAAA 14

RESULT 9  
CN264982  
LOCUS  
DEFINITION  
17000599934645 GRN\_PRENEU Homo sapiens cDNA 5', mRNA sequence.  
ACCSSION  
CN264982  
VERSION  
CN264982.1 GI:47281396  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 469)  
Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J.,  
Li,Y., Xu,C., Pang,R., Guegler,K., Rao,M.S., Mandalam,R.,  
Lebkowski,J and Stanton,L.W.  
Transcriptome characterization elucidates signaling networks that  
control human ES cell growth and differentiation  
Nat. Biotechnol. 22 (6), 707-716 (2004)  
15146197  
Contact: Brandenberger R  
Regenerative Medicine  
Geron Corporation  
230 Constitution Drive, Menlo Park, CA 94025, USA  
Tel: 650 473 8658

Fax: 650 473 7760  
Email: rbrandenberger@ericon.com  
Insert Length: 469 Std Error: 0.00.  
Location/Qualifiers

## FEATURES

Source  
1. .469  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/tissue\_type="embryonic stem cell, retinoic acid and  
mitogen-treated hES cell line H7"  
/clone\_lib="GRN PRENEU"  
/notes="oligo dt primed, full-length enriched cDNA library  
from hES cell line H7 (p29) maintained in feeder-free  
conditions. Embryoid bodies were generated in the presence  
of all-trans retinoic acid and mitogens."

## ORIGIN

Query Match 99.5%; Score 321.4; DB 7; Length 469;  
Best Local Similarity 99.7%; Pred. No. 1.3e-81;  
Matches 322; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CGGGGCCAGGCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCGGG 60  
DB 142 CGGGGCCAGGCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCGGG 201  
QY 61 CATCCTCATGACCATTCGATTCACCTGTTACTGGTATTGGCTCTGTGACAGTCTACCA 120  
DB 202 CATCCTCATGACCATTCGATTCATCCTGTTACTGGTATTGGCTCTGTGACAGTCTACCA 261  
QY 121 TATTATGTTACAGATAATACAGAAAAACGGGGTTACTAGTAGCCGCCCATAGCTGCAA 180  
DB 262 TATTATGTTACAGATAATACAGAAAAACGGGGTTACTAGTAGCCGCCCATAGCTGCAA 321  
QY 181 CTTTGGCACTCCACTGTGCAATGCTGGCCCTGCAAGCTGGGGCTGTTGCCCTGCCCCCT 240  
DB 322 CTTTGGCACTCCACTGTGCAATGCTGGCCCTGCAAGCTGGGGCTGTTGCCCTGCCCCCT 381  
QY 241 TGGTCTGCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 300  
DB 382 TGGTCTGCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 441  
QY 301 AAGTGCACGTGCTTGTGAAAAA 323  
DB 442 AAGTGCACGTGCTTGTGAAAAA 464

RESULT 10  
AW574932/c  
LOCUS  
DEFINITION  
UI-HF-BLO-abp-g-08-0-UI.s1 NIH\_MGC\_37 Homo sapiens cDNA clone  
IMAGE:3057590 3', mRNA sequence.  
AW574932  
AW574932.1 GI:7246471  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.

REFERENCE  
1 (bases 1 to 498)  
NIH-MGC http://mgi.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabs-remail.nih.gov  
The sequence contained an oligo-dT track that was present in the  
oligonucleotide that was used to prime the synthesis of first  
strand cDNA and therefore this may represent a bonafide poly A  
tail. Tissue Procurement: Louis M. Staudt, M.D., Ph.D.  
cDNA Library Preparation: M.B. Soares Lab  
CDNA Library Arrayed by: M.B. Soares Lab  
DNA Sequencing by: M.B. Soares Lab  
Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:  
www-bio.llnl.gov/bbrp/image/image.html  
Seq primer: M13 Forward  
POLYA=Yes.

## FEATURES

Source  
1. .498  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone\_lib="IMAGE:3057590"  
/tissue\_type="lymph"  
/cell\_type="germinal center B cells"  
/cell\_line="MGC85"  
/lab\_host="DH10B (LTI)"  
/clone\_lib="NIH\_MGC\_37"  
/notes="Vector: pT7T3-Pac; Site 1: NotI; Site 2: Eco RI;  
Constructed from size fractionated cytoplasmic mRNA  
(1.5-2.5kb). Directionally cloned. Cells provided by Louis  
M. Staudt, Ph.D. Library preparation by Maria de Fatima  
Bonaldo, Ph.D. and M. Bento Soares, Ph.D."

## ORIGIN

Query Match 99.5%; Score 321.4; DB 1; Length 498;  
Best Local Similarity 99.7%; Pred. No. 1.4e-81;  
Matches 322; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CGGGGCCAGGCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCGGG 60  
DB 335 CGGGGCCAGGCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCGGG 276  
QY 61 CATCCTCATGACCATTCGATTCACCTGTTACTGGTATTGGCTCTGTGACAGTCTACCA 120  
DB 275 CATCCTCATGACCATTCGATTCATCCTGTTACTGGTATTGGCTCTGTGACAGTCTACCA 216  
QY 121 TATTATGTTACAGATAATACAGAAAAACGGGGTTACTAGTAGCCGCCCATAGCTGCAA 180  
DB 215 TATTATGTTACAGATAATACAGAAAAACGGGGTTACTAGTAGCCGCCCATAGCTGCAA 156  
QY 181 CTTTGGCACTCCACTGTGCAATGCTGGCCCTGCAAGCTGGGGCTGTTGCCCTGCCCCCT 240  
DB 155 CTTTGGCACTCCACTGTGCAATGCTGGCCCTGCAAGCTGGGGCTGTTGCCCTGCCCCCT 96  
QY 241 TGGTCTGCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 300  
DB 95 TGGTCTGCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 36  
QY 301 AAGTGCACGTGCTTGTGAAAAA 323  
DB 35 AAGTGCACGTGCTTGTGAAAAA 13

RESULT 11  
CA420737/c

LOCUS  
DEFINITION  
UI-H-FG0-bc-q-b-23-0-UI.s1 NCI CGAP EN1 2 Homo sapiens cDNA clone  
UI-H-FG0-bc-q-b-23-0-UI 3', mRNA sequence.

CA420737  
CA420737.1 GI:24783392  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens (human)  
Homo sapiens

REFERENCE  
1 (bases 1 to 498)  
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabs-remail.nih.gov  
Tissue Procurement: James Martin  
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa





```

Qy 301 AAGTGCACGTGCTTGTAAGAAA 323
    |||||
Db 36 AAGTGCACGTGCTTGTAAGAAA 14

RESULT 13
BM672499/c
LOCUS
DEFINITION
    BM672499 505 bp mRNA linear EST 27-FEB-2002
    UI-E-CQ1-ada-d-12-0-UI.s1 UI-E-CQ1 Homo sapiens cDNA clone
    UI-E-CQ1-ada-d-12-0-UI 3', mRNA sequence.
ACCESSION
    BM672499
VERSION
    BM672499.1 GI:18982397
KEYWORDS
    EST.
SOURCE
    Homo sapiens (human)
ORGANISM
    Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
    Hominidae; Homo.
REFERENCE
    1 (bases 1 to 505)
    Bernaldo,M.F., Lennon,G. and Soares,M.B.
    Normalization and subtraction: two approaches to facilitate gene
    discovery
JOURNAL
    Genome Res. 6 (9), 791-806 (1996)
PUBMED
    8889548
COMMENT
    Contact: Soares, MB
    Coordinated Laboratory for Computational Genomics
    University of Iowa
    375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA
    Tel: 319 335 8250
    Fax: 319 335 9565
    Email: bento-soares@uiowa.edu
    Tissue Procurement: Dr. Gregg Hageman
    cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
    cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
    DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
    Clone Distribution: Researchers may obtain clones from Research
    Genetics (www.resgen.com).
    Seq primer: M13 Forward
    POLYA=Yes.

FEATURES
    source
        1. 505
            /organisms="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="UI-E-CQ1-ada-d-12-0-UI"
            /dev_stage="adult"
            /tissue_type="optic nerve"
            /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
            /clone_lib="UI-E-CQ1"
            /note="Organ: eye; Vector: pTT73-Pac (Pharmacia) with a
            modified polylinker; Site_1: EcoR I; Site_2: Not I;
            UI-E-CQ1 is a normalized cDNA library containing the
            following tissue(s): optic nerve. The library was
            constructed according to Bernaldo, Lennon and Soares,
            Genome Research, 6:791-806, 1996. First strand cDNA
            synthesis was primed with an oligo-dT primer containing a
            Not I site. Double stranded cDNA was ligated to an EcoR I
            adaptor, digested with Not I, and cloned directionally
            into pTT73-Pac vector. The oligonucleotide used to prime
            the synthesis of first-strand cDNA contains a library tag
            sequence that is located between the Not I site and the
            (dT)18 tail. The sequence tag for this library is
            CCATTAGTG. This library was created for the program, Gene
            Discovery in the Visual System, supported by National Eye
            Institute (NEI).
            TAG_TISSUE=human optic nerve
            TAG_LIB=UI-E-CQ1
            TAG_SEQ=CCATTAGTG"

ORIGIN
    Query Match 99.5%; Score 321.4; DB 3; Length 505;
    Best Local Similarity 99.7%; Pred. No. 1.4e-81;
    Matches 322; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGGCCAGGCGCTATGCTCCACCGCAGTGCCTGAACATCTGGGCCCTGATTCCTGGG 60
    |||||
Db 336 CGGGCCAGGCGCTATGCTCCACCGCAGTGCCTGAACATCTGGGCCCTGATTCCTGGG 277

Qy 61 CATCTCATGACCATTTGGATTACCCCTGTGTACTGGTATTCCGGCTCTGTGACAGTCTACCA 120
    |||||
Db 276 CATCTCATGACCATTTGGATTACCTCTGTGTACTGGTATTCCGGCTCTGTGACAGTCTACCA 217

Qy 121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCCGCCCATAGCTGCA 180
    |||||
Db 216 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCCGCCCATAGCTGCA 157

Qy 181 CTTTGCACCTCCACTGTGCAATCTGCGCCTGACCGCTGGGGTGTGGCCCTGCCCCCT 240
    |||||
Db 156 CTTTGCACCTCCACTGTGCAATCTGCGCCTGACCGCTGGGGTGTGGCCCTGCCCCCT 97

Qy 241 TGGTCTCTGCCCCCTAGATACAGCAGTTTATATACCCACACACCTGTCTACAGTGTCTTCAAT 300
    |||||
Db 96 TGGTCTCTGCCCCCTAGATACAGCAGTTTATATACCCACACACCTGTCTACAGTGTCTTCAAT 37

Qy 301 AAGTGCACGTGCTTGTAAGAAA 323
    |||||
Db 36 AAGTGCACGTGCTTGTAAGAAA 14

RESULT 14
BM67351
LOCUS
DEFINITION
    BM67351 510 bp mRNA linear EST 04-MAR-2002
    K-EST0049808 S6SNU620s1 Homo sapiens cDNA clone S6SNU620s1-13-C12
    5', mRNA sequence.
ACCESSION
    BM67351
VERSION
    BM67351.1 GI:19096966
KEYWORDS
    EST.
SOURCE
    Homo sapiens (human)
ORGANISM
    Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
    Hominidae; Homo.
REFERENCE
    1 (bases 1 to 510)
    Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
    Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
    Kim,Y.S.
    21C Frontier Korean EST Project 2001
    Unpublished (2002)
    Contact: Kim YS
    Genome Research Center
    Korea Research Institute of Bioscience & Biotechnology
    52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
    Tel: +82-42-860-4470
    Fax: +82-42-860-4409
    Email: yongseung@mail.kribb.re.kr
    Plate: 13 row: C column: 12
    High quality sequence stop: 510.

FEATURES
    source
        1. 510
            /organisms="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="S6SNU620s1-13-C12"
            /sex="g"
            /tissue_type="Ascites"
            /cell_type="Scattering floating"
            /cell_line="SNU-620"
            /lab_host="Top10F"
            /clone_lib="S6SNU620s1"
            /note="Organ: Stomach; Vector: pcNS; Site_1: EcoRI;
            Site_2: NotI. The poly (A) + RNA was dephosphorylated with
            bacterial alkaline phosphatase (BAP) and then decapped
            with tabacco acid pyrophosphatase (TAP). The decapped
            intact mRNA was ligated with DNA-RNA linker including EcoR
            I site by treatment of T4 RNA ligase and the first strand
            cDNA was synthesized from oligo dT-selected mRNA by

```



priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F' by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library. After analyzing and sequencing about 2,000 - 3,000 colonies in original cDNA library, the abundant cDNAs were selected and amplified by PCR reaction using vector region primer including T7 promoter as 5' primer and N(drt)14 as 3' primer. The PCR products were used as template for synthesis of biotinylated single stranded RNA by in vitro transcription reaction. The synthesized RNA probes were hybridized with antisense single stranded cDNAs prepared from original library and incubated with avidin-gel. After removing DNA-RNA hybrids by centrifuge, the subtracted cDNA libraries were constructed by transformation of the remaining DNA into competent cells E. coli Top10F' with electroporation method."

## ORIGIN

```

Query Match      99.5%; Score 321.4; DB 3; Length 510;
Best Local Similarity 99.7%; Pred. No. 1.4e-81;
Matches 322; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGGGCCAGGCTATGCTTCACGGCAAGTGCCTGAACATCTGGGCCCTGATTCGGG 60
Db 182 CGGGGCCAGGCTATGCTTCACGGCAAGTGCCTGAACATCTGGGCCCTGATTCGGG 241

Qy 61 CATCTCATGACCATTTGATTCACCTCTTACTGTATTCGCTCTGACAGCTACCA 120
Db 242 CATCTCATGACCATTTGATTCACCTCTTACTGTATTCGCTCTGACAGCTACCA 301

Qy 121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCCTGCAA 180
Db 302 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCCTGCAA 361

Qy 181 CTTTGGCACTGACATGTCGATGTCGGCCCTGACGCTGGGGCTGTGGCCCTGCCCT 240
Db 362 CTTTGGCACTGACATGTCGATGTCGGCCCTGACGCTGGGGCTGTGGCCCTGCCCT 421

Qy 241 TGGTCTGCCCTAGATACAGATTTATACCCACACCTGTCTACAGTGTCTATCAAT 300
Db 422 TGGTCTGCCCTAGATACAGATTTATACCCACACCTGTCTACAGTGTCTATCAAT 481

Qy 301 AAAGTGCACGTGCTTGTGAAAAA 323
Db 482 AAAGTGCACGTGCTTGTGAAAAA 504

```

## RESULT 15

```

BU783190/c
LOCUS
DEFINITION in04d02.x1 Human insulinoma Homo sapiens cDNA clone IMAGE:612315
3' similar to SW:IN19_HUMAN P13164 INTERFERON-INDUCIBLE PROTEIN
9-27 ; mRNA sequence.

```

## ACCESSION

```

BU783190
VERSION BU783190.1 GI:23827129

```

## KEYWORDS

```

EST.
Homo sapiens (human)

```

## ORGANISM

```

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

```

## REFERENCE

```

1 (bases 1 to 522)
Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
Lemishka,I., Scarce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
Hillier,L., Marr,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagarishvili,R.,

```

TITLE  
JOURNAL  
COMMENT

Williams,T., Jackson,Y. and Bowers,Y.  
Endocrine Pancreas Consortium  
Unpublished (2000)  
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue  
Endocrine Pancreas Consortium  
Harvard University, Howard Hughes Medical Institute  
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,  
MA 02138  
Tel: 617-495-1812  
Fax: 617-495-8557  
Email: dmelton@bioh.harvard.edu  
Library was constructed by Dr. J. Ferrer In vivo mass-excised to  
pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington  
University Genome Sequencing Center For information on obtaining a  
clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)  
Seq primer: -40UP from Gibco  
High quality sequence stop: 403.

## FEATURES

## source

```

1..522
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6123315"
/tissue_type="insulinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="Human insulinoma"
/notes="Organ: pancreas; Vector: pBluescript SK-; Site 1:
XhoI; Site 2: EcoRI; Constructed with lambda ZAPII system
(Stratagene) by Dr. J. Ferrer, in vivo mass-excised to
pBluescript SK- by Dr. H. Inoue following the Washington
University protocol
(http://genome.wustl.edu/est/lambda_protocol.shtml).
Please contact Hiroshi Inoue, MD/PhD for further
information on this library (Metabolism Division, Permutt
Laboratory, Washington University School of Medicine, Box
8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this
is a Washington University Pancreas EST project library."

```

## ORIGIN

```

Query Match      99.5%; Score 321.4; DB 5; Length 522;
Best Local Similarity 99.7%; Pred. No. 1.4e-81;
Matches 322; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGGGCCAGGCTATGCTTCACGGCAAGTGCCTGAACATCTGGGCCCTGATTCGGG 60
Db 334 CGGGGCCAGGCTATGCTTCACGGCAAGTGCCTGAACATCTGGGCCCTGATTCGGG 275

Qy 61 CATCTCATGACCATTTGATTCACCTCTTACTGTATTCGGCTCTGTGACAGTCTACCA 120
Db 274 CATCTCATGACCATTTGATTCACCTCTTACTGTATTCGGCTCTGTGACAGTCTACCA 215

Qy 121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCCTGCAA 180
Db 214 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCCTGCAA 155

Qy 181 CTTTGGCACTGACATGTCGATGTCGGCCCTGACGCTGGGGCTGTGGCCCTGCCCT 240
Db 154 CTTTGGCACTGACATGTCGATGTCGGCCCTGACGCTGGGGCTGTGGCCCTGCCCT 95

Qy 241 TGGTCTGCCCTAGATACAGATTTATACCCACACCTGTCTACAGTGTCTATCAAT 300
Db 94 TGGTCTGCCCTAGATACAGATTTATACCCACACCTGTCTACAGTGTCTATCAAT 35

Qy 301 AAAGTGCACGTGCTTGTGAAAAA 323
Db 34 AAAGTGCACGTGCTTGTGAAAAA 12

```

Search completed: December 7, 2005, 09:55:54  
Job time : 3689.9 secs

*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 21:43:03 ; Search time 2600.1 Seconds  
(without alignments)  
6427.437 Million cell updates/sec

Title: US-09-980-046B-7

Perfect score: 294

Sequence: 1 cggcgatcgccggcggttat.....agcgagcacctagacaaaaa 294

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_in.\*

3: gb\_env.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pr.\*

9: gb\_ro.\*

10: gb\_scs.\*

11: gb\_sy.\*

12: gb\_un.\*

13: gb\_vi.\*

14: gb\_htg.\*

15: gb\_pl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	270.8	92.1	1736	8 BC030541	BC030541 Homo sapi
2	270.8	92.1	2271	8 AB010427	AB010427 Homo sapi
3	270.8	92.1	4491	8 HS9808338	BSX48190 Homo sapi
4	269.4	91.6	2369	6 AR064037	AR064037 Sequence
5	269.4	91.6	2741	6 CQ850514	CQ850514 Sequence
6	269.4	91.6	2741	8 AK127665	AK127665 Homo sapi
7	269.4	91.6	2743	8 AB209746	AB209746 Homo sapi
8	269.4	91.6	2825	8 BC000201	BC000201 Homo sapi
9	269.4	91.6	2322	6 CQ717617	CQ717617 Sequence
10	269.4	91.6	2351	8 BC002489	BC002489 Homo sapi
11	269.4	91.6	199450	8 AC005674	AC005674 Homo sapi
12	269.4	91.6	240000	14 AC009528	AC009528 Homo sapi
13	267.8	91.1	2850	8 AF020056	AF020056 Homo sapi
14	267.4	91.0	3075	10 BV179203	BV179203 sqm10429
15	256.4	87.2	2047	8 AB168915	AB168915 Macaca fa
16	181	61.6	2115	8 AK222764	AK222764 Homo sapi
17	159	54.1	2125	8 AK222897	AK222897 Homo sapi
18	152.6	51.9	212722	14 AC148123	AC148123 Otolomur

c	19	134.8	45.9	207389	14	AC148146
	20	111.6	38.0	2856	9	BC049117
c	21	111.6	38.0	104872	9	AC084322
c	22	111.6	38.0	347166	9	AC084070
	23	110.8	37.7	2908	9	AF020055
c	24	108.8	37.0	177293	14	AC148073
	25	101.4	34.5	2844	9	BC085864
c	26	101.4	34.5	110000	14	AC091339_01
c	27	101.4	34.5	110000	14	AC091339_02
	28	93.4	31.8	2878	9	AB084171
c	29	80	27.2	165573	14	AC121673
c	30	80	27.2	174283	14	AC137259
c	31	80	27.2	189838	14	AC121037
	32	41.4	14.1	3297	5	AF020054
c	33	38.2	13.0	253631	14	AC126155
c	34	37.4	12.7	160238	9	AC133502
c	35	37	12.6	224778	14	AC079491
	36	36.6	12.4	254878	14	AC110841
	37	35.8	12.2	124645	8	HUAC004638
	38	35.8	12.2	124998	8	AC026397
	39	35.8	12.2	128877	14	AC026399
	40	35.8	12.2	170691	8	AC044802
	41	35.8	12.2	206102	8	AC027131
c	42	35.4	12.0	12768	9	AF263346
	43	35.4	12.0	112467	14	AC164519
c	44	35	11.9	78034	8	AL358473
c	45	35	11.9	157351	9	AC156940

#### ALIGNMENTS

BC030541 1736 bp mRNA linear PRI 12-OCT-2004  
Homo sapiens WD repeat domain 1, transcript variant 2, mRNA (CDNA clone MGC:40281 IMAGE:4403484), complete cds.

BC030541

BC030541.2 GI:34192109

MGC.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 1736)

Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McSwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butlerfield,Y.S., Krzywicki,M.I., Skalska,U., Smalilov,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 1736)

Director MGC Project.

Direct Submission

Submitted (07-MAY-2002) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer





TITLE Full-length human cDNA  
JOURNAL Patent: EP 147413-A 983 18-AUG-2004;  
Research Association for Biotechnology (JP)

FEATURES  
source Location/Qualifiers  
1. .2741  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 91.6%; Score 269.4; DB 6; Length 2741;  
Best Local Similarity 99.6%; Pred. No. 1.1e-77;  
Matches 270; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 21 TCCATGATGCTCTGTCAGGAGTGACAAATCACTACTGAGGAGCCACCCCGCCT 80  
DB 1803 TCCATGATGCTCTGTCAGGAGTGACAAATCACTACTGAGGAGCCACCCCGCCT 1862

QY 81 CTGGATGGACCGAATCAGGAGTAGAGTTTAACTGACAGCGGAACATGTCATTTCTATT 140  
DB 1863 CTGGATGGACCGAATCAGGAGTAGAGTTTAACTGACAGCGGAACATGTCATTTCTATT 1922

QY 141 TCTGTGACGGCGCCCATGCCCCACCCACCAAGAGCAGGAGGCCAGTCATGAC 200  
DB 1923 TCTGTGACGGCGCCCATGCCCCACCCACCAAGAGCAGGAGGCCAGTCATGAC 1982

QY 201 CCTCGTCTCTGCAGGGTGTCTGTACACGTTCTTCTGAAAGCTTTAGACAGTTCAGTTTG 260  
DB 1983 CCTCGTCTCTGCAGGGTGTCTGTACACGTTCTTCTGAAAGCTTTAGACAGTTCAGTTTG 2042

QY 261 CACATGAAAAATAAAGCGAGCACCCTAGACAA 291  
DB 2043 CACATGAAAAATAAAGCGAGCACCCTAAACAA 2073

RESULT 6  
AK127665  
LOCUS Homo sapiens cDNA FLJ45763 fis, clone N1ESE2000698, highly similar  
DEFINITION to WD-repeat protein 1.  
ACCESSION AK127665  
VERSION AK127665.1 GI:34534673  
KEYWORDS oligo capping; fis (full insert sequence).  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.

REFERENCE  
1  
Oshima, A., Takahashi-Fujii, A., Tanase, T., Imose, N., Takeuchi, K.,  
Arita, M., Mueashino, K., Yuuki, H., Hara, H., Sugiyama, T., Irie, R.,  
Otsuki, T., Sato, H., Wakamatsu, A., Umezawa, A., Imabayashi, H.,  
Fukuma, M., Izawa, Y., Hata, J., Ishii, S., Yamamoto, J., Isono, Y.,  
Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H.,  
Matsumoto, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K.,  
Wagatsuma, M., Murakawa, K., Kanehori, K., Sugiyama, A., Kawakami, B.,  
Suzuki, Y., Sugano, S., Nagahara, K., Masuno, Y., Nagai, K. and  
Icogai, T.  
NEDO human cDNA sequencing project  
Unpublished  
2 (bases 1 to 2741)  
Icogai, T. and Yamamoto, J.  
Direct Submission  
Submitted (15-JUL-2003) Takao Iicogai, FLJ Project (HRI Team); 2-6-7  
Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan  
(E-mail: genomic@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)

TITLE NEDO human cDNA sequencing project supported by Ministry of  
Economy, Trade and Industry of Japan; cDNA full insert sequencing:  
Research Association for Biotechnology (RAB); cDNA library  
construction: Helix Research Institute (HRI) (supported by Japan  
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,  
HRI, and Biotechnology Center, National Institute of Technology and  
Evaluation; clone selection for full insert sequencing: HRI and

FEATURES  
source Location/Qualifiers  
1. .2741  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/cell\_type="mesenchymal stem cells"  
/clone\_lib="N1ESE2"  
/notes="cloning vector: pME18SFL3"

ORIGIN  
Query Match 91.6%; Score 269.4; DB 8; Length 2741;  
Best Local Similarity 99.6%; Pred. No. 1.1e-77;  
Matches 270; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 21 TCCATGATGCTCTGTCAGGAGTGACAAATCACTACTGAGGAGCCACCCCGCCT 80  
DB 1803 TCCATGATGCTCTGTCAGGAGTGACAAATCACTACTGAGGAGCCACCCCGCCT 1862

QY 81 CTGGATGGACCGAATCAGGAGTAGAGTTTAACTGACAGCGGAACATGTCATTTCTATT 140  
DB 1863 CTGGATGGACCGAATCAGGAGTAGAGTTTAACTGACAGCGGAACATGTCATTTCTATT 1922

QY 141 TCTGTGACGGCGCCCATGCCCCACCCACCAAGAGCAGGAGGCCAGTCATGAC 200  
DB 1923 TCTGTGACGGCGCCCATGCCCCACCCACCAAGAGCAGGAGGCCAGTCATGAC 1982

QY 201 CCTCGTCTCTGCAGGGTGTCTGTACACGTTCTTCTGAAAGCTTTAGACAGTTCAGTTTG 260  
DB 1983 CCTCGTCTCTGCAGGGTGTCTGTACACGTTCTTCTGAAAGCTTTAGACAGTTCAGTTTG 2042

QY 261 CACATGAAAAATAAAGCGAGCACCCTAGACAA 291  
DB 2043 CACATGAAAAATAAAGCGAGCACCCTAAACAA 2073

RESULT 7  
AB209746  
LOCUS Homo sapiens mRNA for WD repeat-containing protein 1 isoform 1  
DEFINITION variant protein.  
ACCESSION AB209746  
VERSION AB209746.1 GI:62089077  
KEYWORDS FLJ CDNA.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.

REFERENCE  
1  
Totoki, Y., Toyoda, A., Takeda, T., Sakaki, Y., Tanaka, A., Yokoyama, S.,  
Ohara, O., Nagase, T. and F. Kikuno, R.  
None Title  
Published Only in Database (2005)  
2 (bases 1 to 2743)  
Totoki, Y., Toyoda, A., Takeda, T., Sakaki, Y., Tanaka, A., Yokoyama, S.,  
Ohara, O., Nagase, T. and F. Kikuno, R.  
Direct Submission  
Submitted (22-MAR-2005) Osamu Ohara, Kazusa DNA Research Institute,  
Department of Human Gene Research; 2-6-7 Kazusa-kamatari, Kisarazu,  
Chiba, 292-0818, Japan (E-mail: cdnainfo@kazusa.or.jp,  
URL: http://protein.gsc.riken.go.jp/, Tel: 81-438-52-3930,  
Fax: 81-438-52-3931)  
This work was supported in part by the National Project on Protein  
Structural and Functional Analysis, Ministry of Education, Culture  
Sports, Science and Technology of Japan. Totoki Y., Toyoda A.,  
Takeda T., Sakaki Y., Tanaka A., Yokoyama S. RIKEN Genomic Sciences  
Center, 1-7-22 Suehiro-cho, Tsurumi, Yokohama 230-0045, Japan.  
e-mail: aktanaka@postman.riken.go.jp  
URL: http://protein.gsc.riken.go.jp/.

FEATURES  
source Location/Qualifiers  
1. .2741  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/cell\_type="mesenchymal stem cells"  
/clone\_lib="N1ESE2"  
/notes="cloning vector: pME18SFL3"

ORIGIN  
Query Match 91.6%; Score 269.4; DB 8; Length 2741;  
Best Local Similarity 99.6%; Pred. No. 1.1e-77;  
Matches 270; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 21 TCCATGATGCTCTGTCAGGAGTGACAAATCACTACTGAGGAGCCACCCCGCCT 80  
DB 1803 TCCATGATGCTCTGTCAGGAGTGACAAATCACTACTGAGGAGCCACCCCGCCT 1862

QY 81 CTGGATGGACCGAATCAGGAGTAGAGTTTAACTGACAGCGGAACATGTCATTTCTATT 140  
DB 1863 CTGGATGGACCGAATCAGGAGTAGAGTTTAACTGACAGCGGAACATGTCATTTCTATT 1922

QY 141 TCTGTGACGGCGCCCATGCCCCACCCACCAAGAGCAGGAGGCCAGTCATGAC 200  
DB 1923 TCTGTGACGGCGCCCATGCCCCACCCACCAAGAGCAGGAGGCCAGTCATGAC 1982

QY 201 CCTCGTCTCTGCAGGGTGTCTGTACACGTTCTTCTGAAAGCTTTAGACAGTTCAGTTTG 260  
DB 1983 CCTCGTCTCTGCAGGGTGTCTGTACACGTTCTTCTGAAAGCTTTAGACAGTTCAGTTTG 2042

QY 261 CACATGAAAAATAAAGCGAGCACCCTAGACAA 291  
DB 2043 CACATGAAAAATAAAGCGAGCACCCTAAACAA 2073

source	1. .2743 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="ek00180" /tissue_type="aorta endothelial cell" /note="This clone is also named as hek0030002109 vector:pBlueScriptII SK plus" <1. .2743 /gene="WD repeat-containing protein 1 isoform 1 variant" <1. .1977 /gene="WD repeat-containing protein 1 isoform 1 variant" /note="Start codon is not identified." /codon_start=3 /evidence=not experimental /product="WD Repeat-containing protein 1 isoform 1 variant" /protein_id="BAD92993.1" /db_xref="GI:62089078" /translation="RAGPGRRARRGNKWRMPVEIKKVPASLPQVERGVSKIIGD PKGNFLVTNGKCVILNRINDPALDIYTEHAQVWVAKYAPSGFYIASGDVSGKLRI WDTQKEHLAYEQPPAGKIKDIANTEDSKELAVGCEGREKFGAVFLWDSVSGEITGHNVINSVDIKSGRPLRATGSDNCAAFPEGPPKFKFTGVDHFRVNCVRFSPD GNRATASADGOIYIYDGTGEKVGALGSKAHGGIYAIWSPDSTHLLSASGDKTS KIWDSVNSVSTPMPGSDVLDQGLQWQHLHLSVLSGYINLYDRNPFSGELHVI KGHKSIOCLTVHNGKSYIYSGSDGHINWDSYTCENDSPAGKGTNVCISQIIVLL BSGOLISGMDTVRYTSLMRDYSGQGVVLDVQPKCVAVGPGGYAVVVCIGQIVLL KDQRKPSIDNPGEVPEVAVHGGDTVAIGVDGNVRLYSILCTTLKDEGKLEAKG PVTDAVSHGAGFLAVCDASKVTVTVFSVADGTSNNVFGHHKIVCLAWSFDNEHFA SGGMDMMVYVWTLSDPETRVRKIQAHLRLHVSSLAWLDEHTLVTTSHDSVKWETITY "
gene	
CDS	
ORIGIN	Query Match 91.6%; Score 269.4; DB 8; Length 2743; Best Local Similarity 99.6%; Pred. No. 1.1e-77; Matches 270; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy	21 TCCCATGATGCTCTGTCAAGAGTGGACAATCACTACTGAGGAGCCCCACCCCGCCT 80 
Db	1836 TCCCATGATGCTCTGTCAAGAGTGGACAATCACTACTGAGGAGCCCCACCCCGCCT 1895 
Qy	81 CTGGATGACCGGAATCAGGACTAGAGTTTAACTCAGCGGAACATGTCATTTCTTATT 140 
Db	1896 CTGGATGACCGGAATCAGGACTAGAGTTTAACTCAGCGGAACATGTCATTTCTTATT 1955 
Qy	141 TCTGTGACGCGCCCCCATGCCCCACCCACACAGAGGAGGAGGCCCGCCAGTCATGAC 200 
Db	1956 TCTGTGACGCGCCCCCATGCCCCACCCACACAGAGGAGGAGGCCCGCCAGTCATGAC 2015 
Qy	201 CCTGCTCTCTGAGGAGTCTGTACACGTTCTTCTGAAAGCTTTAGACAGTAACAGATTG 260 
Db	2016 CCTGCTCTCTGAGGAGTCTGTACACGTTCTTCTGAAAGCTTTAGACAGTAACAGATTG 2075 
Qy	261 CACATGAAAAATAAAGCGAGCACCTAGACAA 291 
Db	2076 CACATGAAAAATAAAGCGAGCACCTAAACAA 2106 
RESULT 8	
BC000201	
LOCUS	BC000201 2825 bp mRNA linear PRI 29-JUN-2004
DEFINITION	Homo sapiens WD repeat domain 1, transcript variant 1, mRNA (cdna clone MGC:3118 IMAGE:3351638), complete cds.
ACCESSION	BC000201
VERSION	BC000201.1 GI:12652890
KEYWORDS	MGC.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	1 (bases 1 to 2825)
AUTHORS	Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,

TITLE	Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Bueto,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Schetz,T.E., Brownstein,M.J., Ustin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Faney,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywineki,M.I., Skaleka,U., Smailus,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REFERENCE	NIH-MGC Project URL: http://mgc.nci.nih.gov
AUTHORS	Contact: MGC help desk
TITLE	Email: cgapsb-r@mail.nih.gov
JOURNAL	Tissue Procurement: ATCC
REFERENCE	cdna Library Preparation: Rubin Laboratory
AUTHORS	cdna Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
TITLE	DNA Sequencing by: Genome Sequence Centre,
JOURNAL	BC Cancer Agency, Vancouver, BC, Canada
REFERENCE	info@bcgsc.bc.ca
AUTHORS	Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield, Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao, Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin, Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prabh, Parvaneh Saeedi, JR Santos, Angelique Schnerch, Ursula Skalska, Duane Smailus, Jeff Stott, Miranda Tsai, George Yang, Jacque Schein, Asim Siddiqui, Rob Holt, Marco Marra.
REMARK	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
COMMENT	Series: IRAL Plate: 6 Row: k Column: 9 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Similarity but not identity to protein.
FEATURES	Location/Qualifiers
source	1. .2825 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="MGC:3118 IMAGE:3351638" /tissue_type="Eye, retinoblastoma" /clone_lib="NIH MGC 16" /lab_host="DH10B-R" /note="Vector: pOTB7" 1. .2825 /gene="WDRI" /note="synonym: NORI-1" /db_xref="GeneID:9948" /db_xref="MIM:604734" 83_1903 /gene="WDRI" /codon_start=1 /product="WD repeat-containing protein 1, isoform 1" /protein_id="AAH00201.1" /db_xref="GI:12652891"
gene	
CDS	









Qy	261	CACATGAAAAATAAAGCGAGCCTAGACAA	291
Db	2131	CACATGAAAAATAAAGCGAGCACCTAAACAA	2161
RESULT 14			
BV179203/c			
LOCUS			
DEFINITION			
ACCESSION	BV179203	3075 bp	DNA linear STS 10-JUN-2004
VERSION	BV179203.1		squml04292 Human DNA (Sequenom) Homo sapiens STS genomic, sequence tagged site.
KEYWORDS	BV179203.1	GI:48015861	
SOURCE	STS.		
ORGANISM	Homo sapiens (human)		
	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;		
	Hominidae; Homo.		
REFERENCE	1 (bases 1 to 3075)		
AUTHORS	Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M., Cantor,C.R. and Braun,A.		
TITLE	Large-Scale Validation of Single Nucleotide Polymorphisms in Gene Regions		
JOURNAL	Genome Res. (2004) In press		
COMMENT	Contact: Andreas Braun Pharmaceuticals division Seqenom, Inc. 3595 John Hopkins Court, San Diego, CA 92121, USA Tel: 18582029018 Fax: 18582029020 Email: abraun@sequenom.com Primer A: No primer sequence submitted Primer B: No primer sequence submitted STS size: 3075		
FEATURES			
source	Location/Qualifiers		
	1..3075		
	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
	/clone_lib="Human DNA (Sequenom)"		
	<1..>3075		
STS			
ORIGIN			
Query Match	91.0%;	Score 267.4;	DB 10;
Best Local Similarity	98.9%;	Pred. No. 4.9e-77;	Length 3075;
Matches 288;	Conservative	0;	Mismatches 3;
		Indels	0;
Gaps			0;
Qy	21	TCCCATGATGCCTCTGTCAAGGAGTGGACAATCACCTACTGAGGAGCCCCACCCC	80
		TCCTCTGTAAGGAGTGGAACATGTCATGTTTAACTGACGGGAACATGTCTATT	140
Db	1082	TCCCATGATGCCTCTGTCAAGGAGTGGACAATCACCTACTGAGGAGCCCCACCCC	1023
Qy	81	CTGGATGGACCGAATCAGGGACTAGAGTTTAACTGACGGGAACATGTCTATT	140
Db	1022	CTGGATGGACCGAATCAGGGACTAGAGTTTAACTGACGGGAACATGTCTATT	963
Qy	141	TCGTGACCGCCCCCATGCCACCACCAAGAGAGCGAGGCGCCAGTCATGAC	200
Db	962	TCGTGACCGCCCCCATGCCACCACCAAGAGAGCGAGGCGCCAGTCATGAC	903
Qy	201	CCTCGTCTCTGCAAGGTGTCGTACACGTTCTTTCTGAAAAGCTTTAGACAGTAA	260
Db	902	CCTNGTCTCTGCAAGGTGTCGTACACGTTCTTTCTGAAAAGCTTTAGACAGTAA	843
Qy	261	CACATGAAAAATAAAGCGAGCCTTAGACAA	291
Db	842	CACATGAAAAATAAAGCGAGCNCCTTAACAA	812
RESULT 15			
AB168915			
LOCUS			
DEFINITION			
AB168915	Macaca fascicularis testis cDNA, clone: QtsA-15761,	mRNA linear	PRI 18-JUN-2005
similar to			



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 11:31:40 ; Search time 732.068 Seconds  
(without alignments)  
3320.997 Million cell updates/sec

Title: US-09-980-046B-7  
Perfect score: 294  
Sequence: 1 cggcgatgcggcggttat.....agcgagcacctagacaaaaa 294

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_NA Main:  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:  
2: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:  
3: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:  
4: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:  
5: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:  
6: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:  
7: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:  
8: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:  
9: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:  
10: /cgn2\_6/ptodata/1/pubpna/US11\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	270.8	92.1	2271	9 US-10-756-149-1200 Sequence 1200, Ap
2	269.4	91.6	3047	9 US-10-450-763-16835 Sequence 16835, A
3	267.8	91.1	603	5 US-10-060-036-44 Sequence 44, Appl
4	253.6	86.3	452	3 US-09-918-995-35013 Sequence 35013, A
5	244.4	83.1	485	3 US-09-918-995-17083 Sequence 17083, A
6	66.8	22.7	5800	6 US-10-062-674-2202 Sequence 2202, Ap
7	63.2	21.5	506	7 US-10-191-803-483 Sequence 483, App
8	34.2	11.6	598	5 US-10-027-632-321061 Sequence 321061, Sequence 321061,
9	34.2	11.6	598	6 US-10-123-155-464 Sequence 464, App
10	32.8	11.2	941	5 US-10-146-731-464 Sequence 464, App
11	32.8	11.2	941	6 US-10-140-472-464 Sequence 464, App
12	32.8	11.2	941	6 US-10-141-761-464 Sequence 464, App
13	32.8	11.2	941	6 US-10-142-885-464 Sequence 464, App
14	32.8	11.2	941	6 US-10-158-790-464 Sequence 464, App
15	32.8	11.2	941	6 US-10-137-871-464 Sequence 464, App
16	32.8	11.2	941	6 US-10-140-923-464 Sequence 464, App
17	32.8	11.2	941	6 US-10-141-756-464 Sequence 464, App
18	32.8	11.2	941	6 US-10-141-753-464 Sequence 464, App
19	32.8	11.2	941	6 US-10-140-805-464 Sequence 464, App
20	32.8	11.2	941	6 US-10-140-864-464 Sequence 464, App
21	32.8	11.2	941	6 US-10-787-284-2 Sequence 2, Appli
22	32.4	11.0	1559	8 US-09-938-842A-344 Sequence 344, App
23	32.4	11.0	2724	3 US-09-938-842A-344 Sequence 344, App

ALIGNMENTS

RESULT 1

US-10-756-149-1200 2724 3 US-09-938-842A-344  
; Sequence 1200, Application US/10756149  
; Publication No. US20050181375A1  
; GENERAL INFORMATION:  
; APPLICANT: Aziz, Natasha  
; APPLICANT: Zlotnik, Albert  
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND METHODS OF SCREENING FOR MODULATORS OF METASTATIC CANCER  
; TITLE OF INVENTION: METHODS OF SCREENING FOR MODULATORS OF METASTATIC CANCER  
; FILE REFERENCE: file  
; CURRENT APPLICATION NUMBER: US/10756.149  
; CURRENT FILING DATE: 2004-01-12  
; NUMBER OF SEQ ID NOS: 5818  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 1200  
; LENGTH: 2271  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
US-10-756-149-1200

Query Match	92.1%	Score	270.8	DB	9	Length	2271
Best Local Similarity	99.3%	Pred. No.	9.9e-87				
Matches	272	Conservative	0	Mismatches	2	Indels	0
Gaps	0						
Qy	21	TCCCATGATGCTCTGTCAAGGAGTGGACAATCACCTACTGAGGAGCCCCACCCCGCT	80				
Db	1982	TCCCATGATGCTCTGTCAAGGAGTGGACAATCACCTACTGAGGAGCCCCACCCCGCT	2041				
Qy	81	CTGGATGACCGAATCAGGAGTAACTGACGGAACATGTCATTCTCTATT	140				
Db	2042	CTGGATGACCGAATCAGGAGTAACTGACGGAACATGTCATTCTCTATT	2101				
Qy	141	TCTGTGACGCGCCCATGTCCTCCACCCACCAAGAGGAGGAGGCGCCCATGTCATGAC	200				
Db	2102	TCTGTGACGCGCCCATGTCCTCCACCCACCAAGAGGAGGAGGCGCCCATGTCATGAC	2161				
Qy	201	CCTCGTCTCTCAGGGTCTGTACGCTTCTTCTTGAAGCTTTTACAGTAACAGTTTG	260				
Db	2162	CCTCGTCTCTCAGGGTCTGTACGCTTCTTCTTGAAGCTTTTACAGTAACAGTTTG	2221				
Qy	261	CACATGAAAAATAAGCGAGCACCTAGACAAAAA	294				
Db	2222	CACATGAAAAATAAGCGAGCACCTAGACAAAAA	2255				

RESULT 2



```
Dbb 369 GCACATGAAAAATAAGCGACGACCTAAACAA 400

RESULT 5
US-09-918-995-17083
; Sequence 17083, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 17083
; LENGTH: 485
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(485)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-17083

Query Match 83.1%; Score 244.4; DB 3; Length 485;
Best Local Similarity 99.6%; Pred. No. 1.9e-77;
Matches 245; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 46 GCACATCACCCTACTGAGGAGCCGCCCGCTCTGGATGGACCGCAATCAGGGACTAG 105
Db 75 GCACATCACCCTACTGAGGAGCCGCCCGCTCTGGATGGACCGCAATCAGGGACTAG 134

Qy 106 AGTTAACTGACGCGGAACATGTCTATTCTTCTGTCGCGCGCCCGCCATGCCCGCA 165
Db 135 AGTTAACTGACGCGGAACATGTCTATTCTTCTGTCGCGCGCCCGCCATGCCCGCA 194

Qy 156 CCCACCCACAGAGCGAGGCGGCCCGCTCATGACCTCGTCTCTGCGAGGGTGTCTGTAC 225
Db 195 CCCACCCACAGAGCGAGGCGGCCCGCTCATGACCTCGTCTCTGCGAGGGTGTCTGTAC 254

Qy 256 AGCTTCTTCTGAAAGCTTTAGACAGTAACAGTTTGCACATGAAAAATAAGCGACACCT 285
Db 255 AGCTTCTTCTGAAAGCTTTAGACAGTAACAGTTTGCACATGAAAAATAAGCGACACCT 314

Qy 286 AGACAA 291
Db 315 AAACAA 320

RESULT 6
US-10-062-674-2202
; Sequence 2202, Application US/10062674
; Publication No. US20040005559A1
; GENERAL INFORMATION:
; APPLICANT: Loring, Jeanne F.; Kaser, Matthew R.
; TITLE OF INVENTION: MARKERS OF NEURONAL DIFFERENTIATION AND MORPHOGENESIS
; FILE REFERENCE: PA-0026-1 CIP
; CURRENT APPLICATION NUMBER: US/10/062,674
; CURRENT FILING DATE: 2002-01-30
; PRIOR APPLICATION NUMBER: US 09/625,102
; PRIOR FILING DATE: 2000-07-24
; NUMBER OF SEQ ID NOS: 2217
; SOFTWARE: PERL Program
; SEQ ID NO 2202
; LENGTH: 5800
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature

; OTHER INFORMATION: Incyte ID No. US20040005559A1 902527.2
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1) ... (5800)
; OTHER INFORMATION: a, t, c, g, or other
US-10-062-674-2202

Query Match 22.7%; Score 66.8; DB 6; Length 5800;
Best Local Similarity 77.4%; Pred. No. 1.1e-12;
Matches 212; Conservative 0; Mismatches 42; Indels 20; Gaps 10;

Qy 39 AAGGAGTGGACAATCACCT--ACTGAGGAGCCCCCACCACCCCGCCT-CTGGATGGACC--GA 93
Db 2253 AGGAGTGGACAATCACCTTACTGAGTGAGCCCCCACCACCCCGCTACTGGATGGACCTGAA 2312

Qy 94 ATCAGGGACTAGAGTTTAACTGCA--GCGGAACATGTCTTCTCTATT--TCTGTGAGCG 150
Db 2313 TCAGGGGACTAGAGTTTAACTGCA--GCGGAACATGTCTTCTCTATT--TCTGTGAGCG 2372

Qy 151 GCCCCCATGCCCCCACCACCCACCAAGAG-----GCAAGGAGGCGCCAG-TCATGACCCCTC 204
Db 2373 GCGCCCCAATGCCGCCCCACCCACCAAGAGGCGCAGGAGGCGCCAGTTTCATGACCCCTC 2432

Qy 205 GTCTCT--GCAGGGTGTCTGTACACGTTCTTCT--GAAAGCTTTTAGACAGTAACAGTTTGC 261
Db 2433 GTCTCTGCGCAGGGGTGTCTGTACACGTTCTTCTTGAAGGCTTTAGCAGAGTAACAGTTGTC 2492

Qy 262 ACATGA---AAAAATAAGGCGAGCACCTAGACAA 292
Db 2493 ACAGTAGAGGAATAAAGCGAGCACCTAAACAA 2526

RESULT 7
US-10-191-803-483/c
; Sequence 483, Application US/10191803
; Publication No. US20040014040A1
; GENERAL INFORMATION:
; APPLICANT: MENDRICK, Donna
; APPLICANT: PORTER, Mark
; APPLICANT: JOHNSON, Kory
; APPLICANT: HIGGS, Brandon
; APPLICANT: CASTLE, Arthur
; APPLICANT: ELASHOFF, Michael
; TITLE OF INVENTION: Cardiotoxin Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5090US
; CURRENT APPLICATION NUMBER: US/10/191,803
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: US 60/303,819
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/305,623
; PRIOR FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: US 60/369,351
; PRIOR FILING DATE: 2002-04-03
; PRIOR APPLICATION NUMBER: US 60/377,611
; PRIOR FILING DATE: 2002-05-06
; NUMBER OF SEQ ID NOS: 1140
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 483
; LENGTH: 506
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20040014040A1 A1009012
US-10-191-803-483

Query Match 21.5%; Score 63.2; DB 7; Length 506;
Best Local Similarity 68.8%; Pred. No. 7.8e-12;
Matches 121; Conservative 0; Mismatches 43; Indels 12; Gaps 2;

Qy 121 GAACATGTCATTTCTCTATTCTTCTGTGAGCGCGCCCGCCCATGCCCCCACCACCCACCAAGAGG 180
Db 504 GCACCTTCTTTTAACTATTCTGTAGACGCTCC-----CCTTCTGCTGCTGCTGGAG 451
```

```
QY 181 CAGGAGGCCAGTCATGACCCCTGCTCTGTCAGGG-----TGCTGTACAGCTTCTTC 234
Db 450 TGGGAGGGCGCTTGGAGTACCCCTCTCTTCAGGGGTCCATGTCTGTACGGCTCTTT 391
QY 235 TGAAGCTTTAGACAGTAACAGTTTGCACATGAAAAATAAAGCGGAGCACCTAGACA 290
Db 390 TGAGAGCTTTAGACAGTAACAGTTTGCACACAGAAATAAAGCGGAGCACTAARACA 335

RESULT 8
US-10-027-632-321061/c
; Sequence 321061, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 321061
; LENGTH: 598
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-321061

Query Match 11.6%; Score 34.2; DB 5; Length 598;
Best Local Similarity 51.7%; Pred. No. 0.27; 73; Indels 0; Gaps 0;
Matches 78; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 76 CGCCTCTGGATGGACCGAATCAGGGACTAGAGTTTAAGTCCAGCGGAAACATGTCTTCT 135
Db 589 CACCCTTGTCTGTGACCAATACAGACGTCCAGTTTAATTACCAGGAGCTGTCTGGCT 530
QY 136 CTATTTCTGTACGCGCCCGATGCCCGCCACCCACCAAGAGGCGAGGCCCCAGTC 195
Db 529 GATTCAATTTGACATAGACGTATGACCCACCTCGTTTACTAGAGAAAGCGGCTCAAGA 470
QY 196 ATGACCTCTGCTCTGAGGGTGTCTGTACA 226
Db 469 GAGACCCATTTCCTTGGCGGCACAGGAAA 439

RESULT 9
US-10-027-632-321061/c
; Sequence 321061, Application US/10027632
; Publication No. US200204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
```

```
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 321061
; LENGTH: 598
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-321061

Query Match 11.6%; Score 34.2; DB 6; Length 598;
Best Local Similarity 51.7%; Pred. No. 0.27; 73; Indels 0; Gaps 0;
Matches 78; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 76 CGCCTCTGGATGGACCGAATCAGGGACTAGAGTTTAAGTCCAGCGGAAACATGTCTTCT 135
Db 589 CACCCTTGTCTGTGACCAATACAGACGTCCAGTTTAATTACCAGGAGCTGTCTGGCT 530
QY 136 CTATTTCTGTACGCGCCCGATGCCCGCCACCCACCAAGAGGCGAGGCCCCAGTC 195
Db 529 GATTCAATTTGACATAGACGTATGACCCACCTCGTTTACTAGAGAAAGCGGCTCAAGA 470
QY 196 ATGACCTCTGCTCTGAGGGTGTCTGTACA 226
Db 469 GAGACCCATTTCCTTGGCGGCACAGGAAA 439

RESULT 10
US-10-123-155-464/c
; Sequence 464, Application US/10123155
; Publication No. US20030068794A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C30
; CURRENT APPLICATION NUMBER: US/10/123,155
; CURRENT FILING DATE: 2002-04-15
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 464
; LENGTH: 941
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-123-155-464

Query Match 11.2%; Score 32.8; DB 5; Length 941;
Best Local Similarity 9.3%; Pred. No. 1;
```







```

; SEQ ID NO 464
; LENGTH: 941
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-158-790-464

```

[illegible]

Search completed: December 6, 2005, 22:57:09  
Job time : 734.068 secs

*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 11:34:11 ; Search time 200.691 Seconds  
(without alignments)  
455.919 Million cell updates/sec

Title: US-09-980-046B-7  
Perfect score: 294  
Sequence: 1 cggcgatcgccgcggttat.....agcgagcacctagacaaaaa 294

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3289935 seqs, 155610033 residues

Total number of hits satisfying chosen parameters: 6579870

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_NA\_New:\*  
1: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq2:\*  
9: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq3:\*  
10: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	46	15.6	633	US-10-750-185-47850	Sequence 47850, A
2	32.2	11.0	1581	US-10-750-185-29187	Sequence 29187, A
3	28.6	9.7	1138	US-10-750-185-38374	Sequence 38374, A
4	28.6	9.7	1664	US-10-750-185-26660	Sequence 26660, A
5	28	9.5	917	US-10-750-185-36774	Sequence 36774, A
6	28	9.5	1007	US-10-750-185-53113	Sequence 53113, A
7	28	9.5	157224	US-11-112-908-51	Sequence 51, Appl
8	28	9.5	170189	US-11-112-908-50	Sequence 50, Appl
9	27.8	9.5	795	US-10-750-185-30589	Sequence 30589, A
10	27.8	9.5	1303	US-10-750-185-57470	Sequence 57470, A
11	27.6	9.4	600	US-10-750-185-3545	Sequence 3545, Ap
12	27.6	9.4	131855	US-11-112-908-29	Sequence 29, Appl
13	27.6	9.4	143389	US-11-112-908-30	Sequence 30, Appl
14	27.6	9.4	166020	US-11-112-908-28	Sequence 28, Appl
15	27.6	9.4	171247	US-11-112-908-27	Sequence 27, Appl
16	27	9.2	1126	US-10-750-185-45111	Sequence 45111, A
17	27	9.2	1233	US-10-750-185-63362	Sequence 63362, A
18	27	9.2	1316	US-10-750-185-57351	Sequence 57351, A
19	27	9.2	2624	US-10-750-185-29398	Sequence 29398, A
20	26.8	9.1	1437	US-10-750-185-32922	Sequence 32922, A
21	26.8	9.1	2024	US-10-750-185-45936	Sequence 45936, A
22	26.8	9.1	2476	US-10-986-501-86	Sequence 86, Appl
23	26.8	9.1	165883	US-11-112-908-18	Sequence 18, Appl

ALIGNMENTS

RESULT 1

US-10-750-185-47850  
; Sequence 47850, Application US/10750185  
; Publication No. US20050260603A1  
; GENERAL INFORMATION:  
; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MM1100-2  
; CURRENT APPLICATION NUMBER: US/10750,185  
; CURRENT FILING DATE: 2003-12-31  
; PRIOR APPLICATION NUMBER: US 60/437,482  
; PRIOR FILING DATE: 2002-12-31  
; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 47850  
; LENGTH: 633  
; TYPE: DNA  
; ORGANISM: Bovine 19866881311131  
US-10-750-185-47850  
Query Match 15.6%; Score 46; DB 6; Length 633;  
Best Local Similarity 57.7%; Pred. No. 8.7e-07;  
Matches 82; Conservative 0; Mismatches 60; Indels 0; Gaps 0;  
Qy 21 TCCCATGTCCTGTCAAGAGTGCACATCACTACTGAGGAGCCACCCCGCT 80  
Db 203 TCGCAGCAGCGCTCTGTCAAAGAGTGGACAATCGCTACTGAGGAGCCCTCACACCA 262  
Qy 81 CTGGATGGACCGAATCAGGAGCTAGAGTTTAACGTGACGCGGAACATGTCATTTCTCTATT 140  
Db 263 ACTGACACCGAATCAGGAGCTGGTGGTGAACGAGCTCCGCTCTGTCTCTCTCTG 322  
Qy 141 TCTGTGACGGGCCCGCCATGCC 162  
Db 323 CCGCGCCCTGCCCGCGCTCCC 344  
RESULT 2  
US-10-750-185-29187/c  
; Sequence 29187, Application US/10750185

```
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 29187
; LENGTH: 1581
; TYPE: DNA
; ORGANISM: Bovine 19866881898685
; US-10-750-185-29187

Query Match 11.0%; Score 32.2; DB 6; Length 1581;
Best Local Similarity 59.1%; Pred. No. 0.12;
Matches 55; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

Qy 140 TTCTGTGACGGCCCCCATGCCCCACCCACCAAGAGGCGAGGGCCCAAGTCATGA 199
Db 769 TGTGTGTGACCAACCCCGGGCCCCCTCCCCCAAGGCGAGTCGAGGGCTAGTTTGG 710

Qy 200 CCCTCGTCTCTGACGGGTGCTGTACAGTTCT 232
Db 709 GCCAGGGCGCGTAGGGATTGGGACAGGGTCT 677

RESULT 3
US-10-750-185-38374
; Sequence 38374, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 38374
; LENGTH: 1138
; TYPE: DNA
; ORGANISM: Bovine 19866880599576
; US-10-750-185-38374

Query Match 9.7%; Score 28.6; DB 6; Length 1138;
Best Local Similarity 57.1%; Pred. No. 2;
Matches 52; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

Qy 128 TCATTTCCTATTCTGTGACGGCCCCCATGCCCCACCCACCAAGAGGCGAGG 187
Db 24 TCATTTCCTTTTATCTGACTACCCCAAGTCTCTGTGCCCCGCCACAGACAGG 83

Qy 188 GCCCAGTCATGACCCCTCGTCTCTGACGGGTG 218
Db 84 GCTGAGTCAATTCGAGGATCTCGAGTCTG 114

; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 26660
; LENGTH: 1664
; TYPE: DNA
; ORGANISM: Bovine 19866881515211
; US-10-750-185-26660

Query Match 9.7%; Score 28.6; DB 6; Length 1664;
Best Local Similarity 57.1%; Pred. No. 2.4;
Matches 52; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

Qy 25 ATGATGCTCTGTCAAGGAGTGACCAATCACCTACTGAGGAGCCGCCCGCTTGG 84
Db 1095 AGGATGCTCAGTTCTTGGGAGGGTGCTGTCATCTTCTTGGGGCACACACTCCAGCCCCAG 1154

Qy 85 ATGGACCGAATCAGGACTAGATTTAACTG 115
Db 1155 TTGCACCGCATGCACGANTAGACTGAACGG 1185

RESULT 5
US-10-750-185-36774/c
; Sequence 36774, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 36774
; LENGTH: 917
; TYPE: DNA
; ORGANISM: Bovine 19866880785228
; US-10-750-185-36774

Query Match 9.5%; Score 28; DB 6; Length 917;
Best Local Similarity 53.7%; Pred. No. 3;
Matches 58; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

Qy 59 CTGAGGAGCCCCACCCCGCTCTGATGAGCCGAATCAGGGACTAGATTTAACTGCAG 118
```

Db 145 CTGCAAAAGCCAGGCGCGTGGTGAAGAGTGAACCTGGGCGCTCTGCCCTGGCCTG 86  
Qy 119 CGGAACATGTCTATTTCTTATTTCTGTGACGCGCCCGCCCATGCCCCAC 166  
Db 85 GTCAAGAGCATTTGGGCGCTGCACTGTGGGCGCCCGAGCACTC 38

## RESULT 6

US-10-750-185-53113/c  
; Sequence 53113, Application US/10750185  
; Publication No. US20050260603A1  
; GENERAL INFORMATION:  
; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MM1100-2  
; CURRENT APPLICATION NUMBER: US/10/750,185  
; CURRENT FILING DATE: 2003-12-31  
; PRIOR APPLICATION NUMBER: US 60/437,482  
; PRIOR FILING DATE: 2002-12-31  
; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 53113  
; LENGTH: 1007  
; TYPE: DNA  
; ORGANISM: Bovine 19866881478112  
US-10-750-185-53113

Query Match 9.5%; Score 28; DB 6; Length 1007;

Best Local Similarity 48.2%; Pred. No. 3.1;  
Matches 79; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

Qy 23 CCATGATGCTCTGTCAAGAGTGACAAATCACTTACTGAGGAGCCCAACCCCGCTCT 82  
Db 814 CCAGAGCTGCTGGCAGCAGCAGTCTTCTTGTAGATGCCCACTCCCTCC 755  
Qy 83 GGATGAGCCGAATCAGGAGTATGATTTAACTGACGGGAAACATGCTATTTCTATTTC 142  
Db 754 GGGTCCCTTATTTCTGGGAGCTGTAATTTAGACAACCTTCCCGCCAGTTGCTATCTG 695  
Qy 143 TGTGAGCGCCCGCATGCCCGCCCAACCCCAACCAAGAGGAGGAG 186  
Db 694 TGGCAGCAGTCACTTCCCTCAGGCGCTGCGCCACTGCCACTAG 651

## RESULT 7

US-11-112-908-51  
; Sequence 51, Application US/11112908  
; Publication No. US20050260659A1  
; GENERAL INFORMATION:  
; APPLICANT: Harris, Cole  
; APPLICANT: Davis, Lisa M.  
; TITLE OF INVENTION: Breast Cancer Biomarkers  
; FILE REFERENCE: 04-164-US  
; CURRENT APPLICATION NUMBER: US/11/112,908  
; CURRENT FILING DATE: 2005-04-22  
; PRIOR APPLICATION NUMBER: US 60/564,758  
; PRIOR FILING DATE: 2004-04-23  
; PRIOR APPLICATION NUMBER: US 60/575,978  
; PRIOR FILING DATE: 2004-06-01  
; PRIOR APPLICATION NUMBER: US 60/631,702  
; PRIOR FILING DATE: 2004-11-30  
; PRIOR APPLICATION NUMBER: US 60/633,826  
; PRIOR FILING DATE: 2004-12-07  
; NUMBER OF SEQ ID NOS: 511  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 51  
; LENGTH: 157224

; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-11-112-908-51

Query Match 9.5%; Score 28; DB 7; Length 157224;  
Best Local Similarity 51.6%; Pred. No. 35;  
Matches 64; Conservative 0; Mismatches 60; Indels 0; Gaps 0;  
Qy 2 GGCATGCGGCGGCTTATTCCCATGATGCTCTGTCAAGGAGTGGACAATCACTACTG 61  
Db 40234 GGTGCTGCTGCTGCTGGACTCTTGTAGTCACTTGGCTGGGAGTTGAAAGTAGCTGAGA 40293  
Qy 62 AGAGCCCCACCCCGGCTCTGGATGGACCGATCAGGACTAGAGTTTAAGTGCAGCGG 121  
Db 40294 ATGTACCTCTCCAACTCTCATCTTTACAGAGGCCCATGGAGGAAACTGACCTGCTAGG 40353  
Qy 122 AACCA 125  
Db 40354 AACCA 40357

## RESULT 8

US-11-112-908-50  
; Sequence 50, Application US/11112908  
; Publication No. US20050260659A1  
; GENERAL INFORMATION:  
; APPLICANT: Harris, Cole  
; APPLICANT: Davis, Lisa M.  
; TITLE OF INVENTION: Breast Cancer Biomarkers  
; FILE REFERENCE: 04-164-US  
; CURRENT APPLICATION NUMBER: US/11/112,908  
; CURRENT FILING DATE: 2005-04-22  
; PRIOR APPLICATION NUMBER: US 60/564,758  
; PRIOR FILING DATE: 2004-04-23  
; PRIOR APPLICATION NUMBER: US 60/575,978  
; PRIOR FILING DATE: 2004-06-01  
; PRIOR APPLICATION NUMBER: US 60/631,702  
; PRIOR FILING DATE: 2004-11-30  
; PRIOR APPLICATION NUMBER: US 60/633,826  
; PRIOR FILING DATE: 2004-12-07  
; NUMBER OF SEQ ID NOS: 511  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 50  
; LENGTH: 170189  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-11-112-908-50

Query Match 9.5%; Score 28; DB 7; Length 170189;

Best Local Similarity 51.6%; Pred. No. 36;  
Matches 64; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

Qy 2 GGCATGCGGCGGCTTATTCCCATGATGCTCTGTCAAGGAGTGGACAATCACTACTG 61  
Db 72994 GGTGCTGCTGCTGCTGGACTCTTGTAGTCACTTGGCTGGGAGTTGAAAGTAGCTGAGA 73053  
Qy 62 AGAGCCCCACCCCGGCTCTGGATGGACCGAATCAGGACTAGAGTTTAAGTGCAGCGG 121  
Db 73054 ATGTACCTCTCCAACTCTCATCTTTACAGAGGCCCATGGAGGAAACTGACCTGCTAGG 73113  
Qy 122 AACCA 125  
Db 73114 AACCA 73117

## RESULT 9

US-10-750-185-30589/c  
; Sequence 30589, Application US/10750185  
; Publication No. US20050260603A1  
; GENERAL INFORMATION:  
; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard

```
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 30589
; LENGTH: 795
; TYPE: DNA
; ORGANISM: Bovine 19866880821049
US-10-750-185-30589

Query Match          9.5%; Score 27.8; DB 6; Length 795;
Best Local Similarity 59.5%; Pred. No. 3.3;
Matches 47; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

Qy 127 GTATTTCTCTATTCTGTGAGCGGCCCCCATGCCCCACCCACCAAGAGGCGAGG 186
Db 263 GTATCCCTCTAGGTCAGGGAGCGCCCCCAGCGCCCACTCTGGGCCACAGACCAGTA 204

Qy 187 GCGCCAGTCATGACCCCTCG 205
Db 203 GCTCCTGTAGGTCAGTGG 185

RESULT 10
US-10-750-185-57470/c
; Sequence 57470, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 57470
; LENGTH: 1303
; TYPE: DNA
; ORGANISM: Bovine 19866881827269
US-10-750-185-57470

Query Match          9.5%; Score 27.8; DB 6; Length 1303;
Best Local Similarity 62.0%; Pred. No. 4.2;
Matches 44; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

Qy 199 ACCCTCTCTCTGAGGGTGCTCTACACGTTCTCTGAAGCTTTAGACAGTAACAGTT 258
Db 536 AGCTTCTCTATACAGTTGCTAGTAACGGCTCTTTGGAAGCTGTAAACAGACGCGCC 477

Qy 259 TGCACATGAA 269
Db 476 AGTCATGAA 466

RESULT 11
US-10-750-185-3545/c
; Sequence 3545, Application US/10750185
```

```
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3545
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Bovine MMBT12586
US-10-750-185-3545

Query Match          9.4%; Score 27.6; DB 6; Length 600;
Best Local Similarity 55.1%; Pred. No. 3.4;
Matches 54; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

Qy 63 GGAGCCCCCCCCCGCTCTGGATGAGCCGAATCAGGACTAGAGTTTAATCGAGCGGA 122
Db 295 GGTACCTGAAGCAAGTCACTCAACTAAGTTAATGAGAGAGTGGGTTTGAATCGAGGTT 236

Qy 123 ACATGTCATTCTCTATTCTGTGACGCGCCCCCATGC 160
Db 235 ACCTGTCATAGCCCATGCTATTCTTTACCCACGC 198

RESULT 12
US-11-112-908-29
; Sequence 29, Application US/11112908
; Publication No. US20050260659A1
; GENERAL INFORMATION:
; APPLICANT: Harris, Cole
; APPLICANT: Davis, Lisa M.
; TITLE OF INVENTION: Breast Cancer Biomarkers
; FILE REFERENCE: 04-164-US
; CURRENT APPLICATION NUMBER: US/11/112,908
; CURRENT FILING DATE: 2005-04-22
; PRIOR APPLICATION NUMBER: US 60/564,758
; PRIOR FILING DATE: 2004-04-23
; PRIOR APPLICATION NUMBER: US 60/575,978
; PRIOR FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US 60/631,702
; PRIOR FILING DATE: 2004-11-30
; PRIOR APPLICATION NUMBER: US 60/633,826
; PRIOR FILING DATE: 2004-12-07
; NUMBER OF SEQ ID NOS: 511
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 29
; LENGTH: 131855
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-112-908-29

Query Match          9.4%; Score 27.6; DB 7; Length 131855;
Best Local Similarity 50.0%; Pred. No. 45;
Matches 69; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

Qy 142 CTGTGAGCGGCCCCCATGCCCCCACCACCAAGAGGAGGGGCCCATGCATCACC 201
Db 93298 CAGTAGCTCGCGCCCATGGTCCCAAGTCTCAGAACTAGGAGGAGTTGCTTAAG 93357

Qy 202 CTCGTCTCTGACGGGTGCTGTGTACAGCTTCTCTGAAAGCTTTAGACAGTAACAGTTGC 261
```



Ds 93358 GTCATTCAGCCTGGTGACAGAGTGAGATCCTGTCTAAATAAAAAAAAAAGGCACTCT 93417  
Qy 262 ACATGAAAAATAAGCGA 279  
Db 93418 GCACTCCAAAAAGGGAGA 93435

RESULT 13  
US-11-112-908-30  
; Sequence 30, Application US/11112908  
; Publication No. US20050260659A1  
; GENERAL INFORMATION:  
; APPLICANT: Harris, Cole  
; APPLICANT: Davis, Lisa M.  
; TITLE OF INVENTION: Breast Cancer Biomarkers  
; FILE REFERENCE: 04-164-US  
; CURRENT APPLICATION NUMBER: US/11/112,908  
; PRIOR FILING DATE: 2005-04-22  
; PRIOR FILING DATE: 2004-04-23  
; PRIOR FILING DATE: 2004-06-01  
; PRIOR FILING DATE: 2004-06-01  
; PRIOR FILING DATE: 2004-11-30  
; PRIOR FILING DATE: 2004-11-30  
; PRIOR FILING DATE: 2004-12-07  
; NUMBER OF SEQ ID NOS: 511  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 30  
; LENGTH: 143389  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-11-112-908-30

Query Match 9.4%; Score 27.6; DB 7; Length 143389;  
Best Local Similarity 50.0%; Pred. No. 47;  
Matches 69; Conservative 0; Mismatches 69; Indels 0; Gaps 0;  
Qy 142 CTGTGACGCGCCCGCCATGCCCGCCACACAGAGGCGAGGCGCCCGCCAGTCATGACC 201  
Db 61071 CAGTAGCTCGCGCCCGCCATGCCCGCCACAGAGGCGAGGCGCCCGCCAGTCATGACC 61130  
Qy 202 CTCGTCTCTGACGGGTGTCTGTACACGTTCTTCTGAAAGCTTTAGACAGTAACAGTTTGC 261  
Db 61131 GTCATTCAGCCTGGTGACAGTGAGATCCTGTCTTAAATAAAAAAAAAAGGCACTCT 61190  
Qy 262 ACATGAAAAATAAGCGA 279  
Db 61191 GCACTCCAAAAAGGGAGA 61208

RESULT 14  
US-11-112-908-28  
; Sequence 28, Application US/11112908  
; Publication No. US20050260659A1  
; GENERAL INFORMATION:  
; APPLICANT: Harris, Cole  
; APPLICANT: Davis, Lisa M.  
; TITLE OF INVENTION: Breast Cancer Biomarkers  
; FILE REFERENCE: 04-164-US  
; CURRENT APPLICATION NUMBER: US/11/112,908  
; PRIOR FILING DATE: 2005-04-22  
; PRIOR FILING DATE: 2004-04-23  
; PRIOR FILING DATE: 2004-06-01  
; PRIOR FILING DATE: 2004-06-01  
; PRIOR FILING DATE: 2004-11-30  
; PRIOR FILING DATE: 2004-11-30  
; PRIOR FILING DATE: 2004-12-07  
; NUMBER OF SEQ ID NOS: 511  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 28

; LENGTH: 166020  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-11-112-908-28

Query Match 9.4%; Score 27.6; DB 7; Length 166020;  
Best Local Similarity 50.0%; Pred. No. 50;  
Matches 69; Conservative 0; Mismatches 69; Indels 0; Gaps 0;  
Qy 142 CTGTGACGCGCCCGCCATGCCCGCCACACAGAGGCGAGGCGCCCGCCAGTCATGACC 201  
Db 107552 CAGTAGCTCGCGCCCGCCATGCCCGCCAGTACTCAGAGACTGAGGTGGGAGGATTGCTTAAG 107611  
Qy 202 CTCGTCTCTGACGGGTGTCTGTACACGTTCTTCTGAAAGCTTTAGACAGTAACAGTTTGC 261  
Db 107612 GTCATTCAGCCTGGTGACAGTGAGATCCTGTCTTAAATAAAAAAAAAAGGCACTCT 107671  
Qy 262 ACATGAAAAATAAGCGA 279  
Db 107672 GCACTCCAAAAAGGGAGA 107689

RESULT 15  
US-11-112-908-27  
; Sequence 27, Application US/11112908  
; Publication No. US20050260659A1  
; GENERAL INFORMATION:  
; APPLICANT: Harris, Cole  
; APPLICANT: Davis, Lisa M.  
; TITLE OF INVENTION: Breast Cancer Biomarkers  
; FILE REFERENCE: 04-164-US  
; CURRENT APPLICATION NUMBER: US/11/112,908  
; CURRENT FILING DATE: 2005-04-22  
; PRIOR APPLICATION NUMBER: US 60/564,758  
; PRIOR FILING DATE: 2004-04-23  
; PRIOR APPLICATION NUMBER: US 60/575,978  
; PRIOR FILING DATE: 2004-06-01  
; PRIOR APPLICATION NUMBER: US 60/631,702  
; PRIOR FILING DATE: 2004-11-30  
; PRIOR APPLICATION NUMBER: US 60/633,826  
; PRIOR FILING DATE: 2004-12-07  
; NUMBER OF SEQ ID NOS: 511  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 27  
; LENGTH: 171247  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-11-112-908-27

Query Match 9.4%; Score 27.6; DB 7; Length 171247;  
Best Local Similarity 50.0%; Pred. No. 51;  
Matches 69; Conservative 0; Mismatches 69; Indels 0; Gaps 0;  
Qy 142 CTGTGACGCGCCCGCCATGCCCGCCACACAGAGGCGAGGCGCCCGCCAGTCATGACC 201  
Db 157511 CAGTAGCTCGCGCCCGCCATGCCCGCCAGTACTCAGAGACTGAGGTGGGAGGATTGCTTAAG 157570  
Qy 202 CTCGTCTCTGACGGGTGTCTGTACACGTTCTTCTGAAAGCTTTAGACAGTAACAGTTTGC 261  
Db 157571 GTCATTCAGCCTGGTGACAGTGAGATCCTGTCTTAAATAAAAAAAAAAGGCACTCT 157630  
Qy 262 ACATGAAAAATAAGCGA 279  
Db 157631 GCACTCCAAAAAGGGAGA 157648

Search completed: December 6, 2005, 23:16:54  
Job time : 202.691 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 21:43:57 ; Search time 3355.79 Seconds  
(without alignments)  
4099.008 Million cell updates/sec

Title: US-09-980-046B-7  
Perfect score: 294  
Sequence: 1 cggcgatcgcgcggttat.....agcgagcactagacaaaaa 294

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST: \*  
1: gb\_est1: \*  
2: gb\_est2: \*  
3: gb\_est3: \*  
4: gb\_hcc: \*  
5: gb\_est4: \*  
6: gb\_est5: \*  
7: gb\_est6: \*  
8: gb\_est7: \*  
9: gb\_ges1: \*  
10: gb\_ges2: \*  
11: gb\_ges3: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	271.4	92.3	670	BM470718	AGENCOURT
C 2	270.8	92.1	330	AA804321	AA804321 nw26911.s
3	270.8	92.1	346	BQ690721	AGENCOURT
C 4	270.8	92.1	397	AA599515	AA599515 ag26e07.s
C 5	270.8	92.1	411	BQ018738	UI-H-DH1-
C 6	270.8	92.1	419	BQ652690	AGENCOURT
C 7	270.8	92.1	470	AA599195	AA599195 ag34a07.s
C 8	270.8	92.1	575	BQ008059	UI-H-E10-
C 9	270.8	92.1	588	BUE32242	UI-H-FE1-
C 10	270.8	92.1	753	CA431219	UI-H-FG1-
C 11	270.4	92.0	428	AA156259	AA156259 zo48b04.r
12	269.4	91.6	543	CR736990	CR736990 CR736990
13	269.4	91.6	628	CRK430066	CRK430066 oJ44f10.y
14	269.4	91.6	628	CNA428169	CNA428169 170005999
15	269.4	91.6	650	CB114667	CB114667 K-EST0158
16	269.4	91.6	692	AV718358	AV718358 AV718358
17	269.4	91.6	723	CR990355	CR990355 CR990355
18	269.4	91.6	740	BG681135	BG681135 602828965
19	269.4	91.6	809	BG169773	BG169773 602324595
20	269.4	91.6	874	CV810559	CV810559 AGENCOURT
21	269.4	91.6	895	CV811638	CV811638 AGENCOURT
22	269.4	91.6	997	BM458195	BM458195 AGENCOURT

23	269.4	91.6	1038	3	BM543106	BM543106 AGENCOURT
24	269.4	91.6	1095	3	BM456796	BM456796 AGENCOURT
25	269.4	91.6	1542	4	CR595687	full-leng
26	269.4	91.6	1761	4	HSM800862	AL110208 Homo sapi
27	269.4	91.6	2785	4	HSM800418	AL050108 Homo sapi
C 28	269.2	91.6	391	3	BM670864	BM670864 UI-E-DW1-
C 29	269.2	91.6	459	3	BM991388	BM991388 UI-H-DH0-
C 30	269.2	91.6	466	6	CA437724	UI-H-DH0-
C 31	269.2	91.6	468	1	AI281760	AI281760 qt86912.x
C 32	269	91.5	1185	1	AL563732	AL563732 AL563732
33	268.8	91.4	899	5	BQ691167	BQ691167 AGENCOURT
34	268.4	91.3	717	7	CT002650	CT002650 CT002650
35	267.8	91.1	488	3	BM832046	BM832046 K-EST0106
36	267.8	91.1	563	3	BM759791	BM759791 K-EST0040
37	267.8	91.1	754	5	BQ431810	BQ431810 AGENCOURT
38	267.8	91.1	1041	5	BX405988	BX405988 BX405988
C 39	267.6	91.0	392	1	AI147712	AI147712 qb43c08.x
C 40	267.6	91.0	1076	1	AL579733	AL579733 AL579733
41	267.4	91.0	368	8	244036	244036 HSC1RB041 n
42	267	90.8	858	5	BU540159	BU540159 AGENCOURT
C 43	266.2	90.5	600	1	AI991096	AI991096 wu38a06.x
C 44	266.2	90.5	968	2	BG697854	BG697854 602661067
C 45	266	90.5	419	1	AA400985	AA400985 zu50e04.s

ALIGNMENTS

RESULT 1  
BM470718  
LOCUS  
DEFINITION AGENCOURT 6463211 NIH\_MGC\_71 Homo sapiens cDNA clone IMAGE:5533535  
5', mRNA sequence.  
ACCESSION BM470718  
VERSION BM470718.1 GI:18519760  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE 1 (bases 1 to 670)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Straubeberg, Ph.D.  
Email: cga@rs-remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM12218 row: i column: 24  
High quality sequence stop: 622.  
Location/Qualifiers  
1. .670  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5533535"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2.1 kb."

Query Match 92.3%; Score 271.4; DB 3; Length 670;  
Best Local Similarity 99.6%; Pred. No. 2.7e-73;  
Matches 272; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 21 TCCCATGATGCTCTGTCAAGGAGTGACAAATCACCTACTGAGGAGCCCCACCCCGCCT 80  
 DB 387 TCCCATGATGCTCTGTCAAGGAGTGACAAATCACCTACTGAGGAGCCCCACCCCGCCT 446  
 QY 81 CTGGATGGACCGAATCAGGAGCTAGAGTTTAACTGCGAGCGGAACATGTCTCTATT 140  
 DB 447 CTGGATGGACCGAATCAGGAGCTAGAGTTTAACTGCGAGCGGAACATGTCTCTATT 506  
 QY 141 TCTGTGACGGCCCCCATGCCCCACCCACCAACAGAGGAGGAGGCCCCAGTCTATGAC 200  
 DB 507 TCTGTGACGGCCCCCATGCCCCACCCACCAACAGAGGAGGAGGCCCCAGTCTATGAC 566  
 QY 201 CCTCGTCTCTGCGAGGGTGTCTGTACACGTTCTTCTGAAAGCTTTAGACAGTAACAGTTTG 260  
 DB 567 CCTCGTCTCTGCGAGGGTGTCTGTACACGTTCTTCTGAAAGCTTTAGACAGTAACAGTTTG 626  
 QY 261 CACATGAAAAATAAAGCGAGCAGCTAGACAAAA 293  
 DB 627 CACATGAAAAATAAAGCGAGCAGCTAAACAAAA 659

RESULT 2  
 AA804321/c  
 LOCUS  
 DEFINITION n426g11.s1 NCI\_CGAP\_GCB0 Homo sapiens cDNA clone IMAGE:1241636 3', mRNA linear EST 18-FEB-1998  
 mRNA sequence.  
 ACCESSION AA804321  
 VERSION AA804321.1 GI:2873608  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 330)  
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
 Unpublished (1997)  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgapsb@mail.nih.gov](mailto:cgapsb@mail.nih.gov)  
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman, Ph.D., Gerald Marti, M.D.  
 CDNA Library Preparation: Louis M. Staudt, M.D., Ph.D.  
 CDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: [www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)  
 Insert Length: 2436 Std Error: 0.00  
 Seq primer: -40m13 fwd. ET from Amersham  
 High quality sequence stop: 228.  
 Location/Qualifiers  
 1..330  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:1241636"  
 /tissue\_type="germinal center B-cells"  
 /lab\_host="SOLR (kanamycin resistant)"  
 /clone\_lib="NCI\_CGAP\_GCB0"  
 /note="Organ: tonsil; Vector: Bluescript SK-; Site 1: EcORI; Site 2: XhoI; Cloned unidirectionally. Primer: Oligo dt. Germinal center B-cells Library constructed by Dr. L. Staudt (NCI). 5' adaptor sequence: 5' GAATTCGGCAGCAG 3' 3' adaptor sequence: 5' CTCGAGTTTCTTTTCTTTTCTTTT 3' Average insert size: 1.1 kb."

FEATURES  
 source  
 1..330  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:1241636"  
 /tissue\_type="germinal center B-cells"  
 /lab\_host="SOLR (kanamycin resistant)"  
 /clone\_lib="NCI\_CGAP\_GCB0"  
 /note="Organ: tonsil; Vector: Bluescript SK-; Site 1: EcORI; Site 2: XhoI; Cloned unidirectionally. Primer: Oligo dt. Germinal center B-cells Library constructed by Dr. L. Staudt (NCI). 5' adaptor sequence: 5' GAATTCGGCAGCAG 3' 3' adaptor sequence: 5' CTCGAGTTTCTTTTCTTTTCTTTT 3' Average insert size: 1.1 kb."

ORIGIN  
 Query Match 92.1%; Score 270.8; DB 1; Length 330;  
 Best Local Similarity 99.3%; Pred. No. 3.6e-73;  
 Matches 272; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 21 TCCCATGATGCTCTGTCAAGGAGTGACAAATCACCTACTGAGGAGCCCCACCCCGCCT 80  
 DB 277 TCCCATGATGCTCTGTCAAGGAGTGACAAATCACCTACTGAGGAGCCCCACCCCGCCT 218  
 QY 81 CTGGATGGACCGAATCAGGAGCTAGAGTTTAACTGCGAGCGGAACATGTCTCTATT 140  
 DB 217 CTGGATGGACCGAATCAGGAGCTAGAGTTTAACTGCGAGCGGAACATGTCTCTATT 158  
 QY 141 TCTGTGACGGCCCCCATGCCCCACCCACCAACAGAGGAGGAGGCCCCAGTCTATGAC 200  
 DB 157 TCTGTGACGGCCCCCATGCCCCACCCACCAACAGAGGAGGAGGCCCCAGTCTATGAC 98  
 QY 201 CCTCGTCTCTGCGAGGGTGTCTGTACACGTTCTTCTGAAAGCTTTAGACAGTAACAGTTTG 260  
 DB 97 CCTCGTCTCTGCGAGGGTGTCTGTACACGTTCTTCTGAAAGCTTTAGACAGTAACAGTTTG 38  
 QY 261 CACATGAAAAATAAAGCGAGCAGCTAGACAAAA 294  
 DB 37 CACATGAAAAATAAAGCGAGCAGCTTAAAAAANA 4

RESULT 3  
 BQ690721  
 LOCUS  
 DEFINITION BQ690721\_8047310 NIH\_MGC\_110 Homo sapiens cDNA clone IMAGE:6207332 5', mRNA sequence.  
 ACCESSION BQ690721  
 VERSION BQ690721.1 GI:21816037  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 346)  
 NIH-MGC <http://mgc.nci.nih.gov/>.  
 National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgapsb@mail.nih.gov](mailto:cgapsb@mail.nih.gov)  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Plate: LLCM2363 row: d column: 21  
 High quality sequence stop: 345.  
 Location/Qualifiers  
 1..346  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6207332"  
 /tissue\_type="ductal carcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH MGC 110"  
 /note="Organ: pancreas; Vector: pOT87; Site 1: XhoI; Site 2: EcORI; cDNA made by oligo-dT priming. Directionally cloned into EcORI/XhoI sites using the following 5' adaptor: GGCAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."

FEATURES  
 source  
 1..346  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6207332"  
 /tissue\_type="ductal carcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH MGC 110"  
 /note="Organ: pancreas; Vector: pOT87; Site 1: XhoI; Site 2: EcORI; cDNA made by oligo-dT priming. Directionally cloned into EcORI/XhoI sites using the following 5' adaptor: GGCAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."

ORIGIN  
 Query Match 92.1%; Score 270.8; DB 5; Length 346;  
 Best Local Similarity 99.3%; Pred. No. 3.6e-73;  
 Matches 272; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 21 TCCCATGATGCTCTGTCAAGAGTGGACAAATCACCTACTGAGGAGCCCCACCCCGCCT 80  
 |||||  
 Db 62 TCCCATGATGCTCTGTCAAGAGTGGACAAATCACCTACTGAGGAGCCCCACCCCGCCT 121  
 |||||  
 QY 81 CTGGATGACCGAATCAGGAGTACAGTTTAACCTGACGAGGAAACATGTCATTCTCTATT 140  
 |||||  
 Db 122 CTGGATGACCGAATCAGGAGTACAGTTTAACCTGACGAGGAAACATGTCATTCTCTATT 181  
 |||||  
 QY 141 TCTGTGACGCGCCCCCATCGCCCCACCCACCAAGAGGAGGAGGCCCGCAGTCATGAC 200  
 |||||  
 Db 182 TCTGTGACGCGCCCCCATCGCCCCACCCACCAAGAGGAGGAGGCCCGCAGTCATGAC 241  
 |||||  
 QY 201 CCTCTCTCTGACGGGTCTGTACACGTTCTTCTGAAAGCTTTAGACAGTAACAGTTTG 260  
 |||||  
 Db 242 CCTCTCTCTGACGGGTCTGTACACGTTCTTCTGAAAGCTTTAGACAGTAACAGTTTG 301  
 |||||  
 QY 261 CACATGAAAAATAAAGCGGACCTAGACAAAA 294  
 |||||  
 Db 302 CACATGAAAAATAAAGCGGACCTAAAAAAA 335  
 |||||

RESULT 4  
 AA599515/c  
 LOCUS  
 DEFINITION  
 IMAGE:1090692 3', mRNA sequence.  
 ACCESSION  
 VERSION  
 AA599515.1 GI:2433140  
 KEYWORDS  
 EST.  
 SOURCE  
 Homo sapiens (human)  
 ORGANISM  
 Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Hominidae; Homo.  
 1 (bases 1 to 397)  
 Jia,L., Robey,P., Young,M., Schuler,G., Powell,J., Yang,L.,  
 Lennon,G., Hillier,L., Allen,M., Bowles,L., Geisel,S., Kucaba,T.,  
 Marra,M., Martin,J., Steptoe,M., Tan,F., Theising,B., Bowers,Y.,  
 Wylie,T., Waterston,R., Wilson,R. and Francomano,C.  
 WashU-MGB/NHGRI EST Project  
 Unpublished (1997)  
 JOURNAL  
 COMMENT  
 Contact: Wilson RK / Jia L  
 WashU-MGB/NHGRI EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 This clone is available royalty-free through LLNL ; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 Insert Length: 723 Std Error: 0.00  
 Seq primer: -40m13 fwd. ET from Amersham  
 High quality sequence stop: 391.  
 Location/Qualifiers

FEATURES  
 source  
 1..397  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:1090692"  
 /sex="mixed"  
 /tissue\_type="bone marrow stroma"  
 /dev\_stage="mixed"  
 /lab\_host="XLI-Blue MRF"/SOLR"  
 /clone\_lib="Human bone marrow stromal cells"  
 /note="Vector: pBluescript; Site\_1: EcoRI; Site\_2: XhoI;  
 mRNA made from human bone marrow stroma, cDNA made by  
 oligo-dT priming. Directionally cloned. Size-selected for  
 average insert size >0.5 kb. Library constructed by Dr.  
 Marian Young and Dr. Pamela Gehron Robey (NIHCR). Library  
 supplied by Dr. Libin Jia (NHGRI)"

ORIGIN

Query Match 92.1%; Score 270.8; DB 1; Length 397;

Best Local Similarity 99.3%; Pred. No. 3.7e-73;  
 Matches 272; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 21 TCCCATGATGCTCTGTCAAGAGTGGACAAATCACCTACTGAGGAGCCCCACCCCGCCT 80  
 |||||  
 Db 275 TCCCATGATGCTCTGTCAAGAGTGGACAAATCACCTACTGAGGAGCCCCACCCCGCCT 216  
 |||||  
 QY 81 CTGGATGACCGAATCAGGAGTACAGTTTAACCTGACGAGGAAACATGTCATTCTCTATT 140  
 |||||  
 Db 215 CTGGATGACCGAATCAGGAGTACAGTTTAACCTGACGAGGAAACATGTCATTCTCTATT 156  
 |||||  
 QY 141 TCTGTGACGCGCCCCCATCGCCCCACCCACCAAGAGGAGGAGGCCCGCAGTCATGAC 200  
 |||||  
 Db 155 TCTGTGACGCGCCCCCATCGCCCCACCCACCAAGAGGAGGAGGCCCGCAGTCATGAC 96  
 |||||  
 QY 201 CCTCTCTCTGACGGGTCTGTACACGTTCTTCTGAAAGCTTTAGACAGTAACAGTTTG 260  
 |||||  
 Db 95 CCTCTCTCTGACGGGTCTGTACACGTTCTTCTGAAAGCTTTAGACAGTAACAGTTTG 36  
 |||||  
 QY 261 CACATGAAAAATAAAGCGGACCTAGACAAAA 294  
 |||||  
 Db 35 CACATGAAAAATAAAGCGGACCTAAAAAAA 2  
 |||||

RESULT 5

BQ018738/c

LOCUS

DEFINITION

IMAGE:5823665 3', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Hominidae; Homo.  
 1 (bases 1 to 411)  
 NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP).  
 Tumor Gene Index  
 Unpublished (1997)  
 JOURNAL  
 COMMENT  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapsb-r@mail.nih.gov  
 Tissue Procurement: Dr. Jose Mercuende  
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov  
 Seq primer: M13 FORWARD  
 POLYA-Yes.  
 Location/Qualifiers

FEATURES

source

1..411

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:5823665"

/tissue\_type="Metastatic Chondrosarcoma"

/dev\_stage="Adult"

/lab\_host="DH10B (Life Technologies)"

/clone\_lib="NCI CGAP DH1"

/note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a  
 modified polylinker; Site\_1: EcoRI; Site\_2: Not I;  
 NCI CGAP DH1 is a normalized cDNA library containing the  
 following tissue(s): VS-8 Cell line from Metastatic  
 Chondrosarcoma in Lung. The library was constructed  
 according to Bonaldo, Lennon and Soares, Genome Research,  
 6:791-806, 1996. First strand cDNA synthesis was primed  
 with an oligo-dT primer containing a Not I site. Double  
 stranded cDNA was ligated to an EcoRI adaptor, digested  
 with Not I, and cloned directionally into pT73-Pac  
 vector. The oligonucleotide used to prime the synthesis of

first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is AGATCATGTC.

TAG\_TISSUE=lung  
TAG\_LIB=UI-H-DH1  
TAG\_SEQ=AGATCATGTC"

## ORIGIN

```

Query Match      92.1%; Score 270.8; DB 3; Length 411;
Best Local Similarity 99.3%; Pred. No. 3.7e-73;
Matches 272; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 21 TCCCATGATGCTCTGTCAAGAGTGGACATACCTACTGAGGAGCCACCCCGCCT 80
    |||
Db 283 TCCCATGATGCTCTGTCAAGAGTGGACATACCTACTGAGGAGCCACCCCGCCT 224
    |||

Qy 81 CTGGATGGACCGAATCAGGACTAGAGTTTAACTGCGAGCGGAACATGTCATTTCTCTATT 140
    |||
Db 223 CTGGATGGACCGAATCAGGACTAGAGTTTAACTGCGAGCGGAACATGTCATTTCTCTATT 164
    |||

Qy 141 TCTGTGAGCGCCGCCATGCCCCACCCACCAAGAGGAGGAGGCGCCAGTCATGAC 200
    |||
Db 163 TCTGTGAGCGCCGCCATGCCCCACCCACCAAGAGGAGGAGGCGCCAGTCATGAC 104
    |||

Qy 201 CCTCGTCTCTGCGAGGGTCTGTACAGCTTCTTCTGAAAGCTTTAGACAGTAAAGTTTG 260
    |||
Db 103 CCTGTCTCTGCGAGGGTCTGTACAGCTTCTTCTGAAAGCTTTAGACAGTAAAGTTTG 44
    |||

Qy 261 CACATGAAAAATAAAGCGAGCAGCCTAGACAAAAA 294
    |||
Db 43 CACATGAAAAATAAAGCGAGCAGCCTAAACAAAAA 10
    |||

```

## RESULT 6

```

BQ652690
LOCUS      BQ652690          419 bp    mRNA    linear    EST 15-JUL-2002
DEFINITION AGENCOURT_8488531 NIH_MGC_100 Homo sapiens cDNA clone IMAGE:6296172
            5', mRNA sequence.
ACCESSION  BQ652690
VERSION    BQ652690.1 GI:21776862
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE  1 (bases 1 to 419)
            NIH-MGC http://mgc.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgabbs-remail.nih.gov
            Tissue Procurement: CGAP (Stanford)
            CDNA Library Preparation: Rubin Laboratory
            CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone Distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2503 row: j column: 13
            High quality sequence stop: 418.
            Location/Qualifiers
                1..419
                    /organism="Homo sapiens"
                    /mol_type="mRNA"
                    /db_xref="taxon:9606"
                    /clone="IMAGE:6296172"
                    /tissue_type="hepatocellular carcinoma, cell line"
                    /lab_host="DH10B (phage-resistant)"
                    /clone_lib="NIH_MGC_100"
                    /note="Organ: liver; Vector: pOTB7; Site 1: XhoI; Site 2:
                    EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                    into EcoRI/XhoI sites using the following 5' adaptor:

```

## FEATURES

```

source
1..419
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6296172"
/tissue_type="hepatocellular carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_100"
/note="Organ: liver; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:

```

## ORIGIN

```

Query Match      92.1%; Score 270.8; DB 5; Length 419;
Best Local Similarity 99.3%; Pred. No. 3.7e-73;
Matches 272; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 21 TCCCATGATGCTCTGTCAAGAGTGGACATACCTACTGAGGAGCCACCCCGCCT 80
    |||
Db 134 TCCCATGATGCTCTGTCAAGAGTGGACATACCTACTGAGGAGCCACCCCGCCT 193
    |||

Qy 81 CTGGATGGACCGAATCAGGACTAGAGTTTAACTGCGAGCGGAACATGTCATTTCTCTATT 140
    |||
Db 194 CTGGATGGACCGAATCAGGACTAGAGTTTAACTGCGAGCGGAACATGTCATTTCTCTATT 253
    |||

Qy 141 TCTGTGAGCGCCGCCATGCCCCACCCACCAAGAGGAGGAGGCGCCAGTCATGAC 200
    |||
Db 254 TCTGTGAGCGCCGCCATGCCCCACCCACCAAGAGGAGGAGGCGCCAGTCATGAC 313
    |||

Qy 201 CCTCGTCTCTGCGAGGGTCTGTACAGCTTCTTCTGAAAGCTTTAGACAGTAAAGTTTG 260
    |||
Db 314 CCTGTCTCTGCGAGGGTCTGTACAGCTTCTTCTGAAAGCTTTAGACAGTAAAGTTTG 373
    |||

Qy 261 CACATGAAAAATAAAGCGAGCAGCCTAGACAAAAA 294
    |||
Db 374 CACATGAAAAATAAAGCGAGCAGCCTAAACAAAAA 407
    |||

```

## RESULT 7

```

AA599195/c
LOCUS      AA599195          470 bp    mRNA    linear    EST 20-JUN-2002
DEFINITION AG34407.81 Human bone marrow stromal cells Homo sapiens cDNA clone
            IMAGE:1091412 3' similar to WP:C04F6.4 CE03924 BETA TRANSDUCIN ;,
            mRNA sequence.
ACCESSION  AA599195
VERSION    AA599195.1 GI:2432820
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE  1 (bases 1 to 470)
            Jia,L., Robey,P., Young,M., Schuler,G., Powell,J., Yang,L.,
            Lennon,G., Hillier,L., Allen,M., Bowles,L., Geisler,S., Kucaba,T.,
            Marra,M., Martin,J., Steptoe,M., Tan,F., Theising,B., Bowers,Y.,
            Wyllie,T., Waterston,R., Wilson,R. and Francimano,C.
            WashU-MGB/NHGRI EST Project
            Unpublished (1997)
            Contact: Wilson RK / Jia L
            WashU-MGB/NHGRI EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Possible reversed clone: similarity on wrong strand
            Insert Length: 2461 Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 321.
            Location/Qualifiers
                1..470
                    /organism="Homo sapiens"
                    /mol_type="mRNA"
                    /db_xref="taxon:9606"
                    /clone="IMAGE:1091412"

```

## FEATURES

```

source
1..470
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1091412"

```

GGCAGCAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."



FEATURES	POLYA=Yes.		JOURNAL COMMENT	Tumor Gene Index Unpublished (1997) Contact: Robert Strausberg, Ph.D. Email: cgabbs-r@mail.nih.gov Tissue Procurement: James Martin cDNA Library preparation: Dr. M. Bento Soares, University of Iowa cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa DNA Sequencing by: Dr. M. Bento Soares, University of Iowa Clone Distribution: Clone distribution information can be obtained from Dr. M. Bento Soares, bento-soares@uiowa.edu Seq primer: M13 FORWARD POLYA=Yes.			
	Location/Qualifiers	Location/Qualifiers					
source	1. .588	1. .753	FEATURES	source			
	/organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="UI-H-FGI-beg-e-07-0-UI" /tissue_type="Cell lines" /dev_stage="Adult" /lab_host="DH10B (Life Technologies)" /clone_lib="NCI CGAP FGI" /notes="Organ: Chondrosarcoma; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP_FGI is a normalized cDNA library derived from a pool of mRNA obtained from 3 cell lines from grade II chondrosarcoma tissues. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CGTACGGAC. The cell lines were provided by Dr James Martin from the University of Iowa. TAG_TISSUE=Human grade 2 chondrosarcoma cell line pool TAG_LIB=UI-H-FGI TAG_SEQ=CGCTACGGAC"						
ORIGIN							
Query Match 92.1%; Score 270.8; DB 5; Length 588; Best Local Similarity 99.3%; Pred. No. 4e-73; Matches 272; Conservative 0; Mismatches 2; Indels 0; Gaps 0;							
Qy	21	TCCCATGATGCTCTGTCAAGGAGTGACAAATCACTACTGAGGAGCCCAACCCCGCCT	80				
Db	283	TCCCATGATGCTCTGTCAAGGAGTGACAAATCACTACTGAGGAGCCCAACCCCGCCT	224				
Qy	81	CTGGATGACCGAATCAGGAGTAGAGTTTAACTGCGAGCGGAACATGCTCTTCTATT	140				
Db	223	CTGGATGACCGAATCAGGAGTAGAGTTTAACTGCGAGCGGAACATGCTCTTCTATT	164				
Qy	141	TCTGTGAGCGGCCCATGCCCCACCCACCAAGAGGAGGAGGCGCCAGTCATGAC	200				
Db	163	TCTGTGAGCGGCCCATGCCCCACCCACCAAGAGGAGGAGGCGCCAGTCATGAC	104				
Qy	201	CCTCGTCTCTGCGGGGTGTGTACAGTCTTCTGAAAGCTTTAGACAGTAACAGTTTG	260				
Db	103	CCTCGTCTCTGCGGGGTGTGTACAGTCTTCTGAAAGCTTTAGACAGTAACAGTTTG	44				
Qy	261	CACATGAAAAATAAGCGGAGCACCTAGACAAAAA	294				
Db	43	CACATGAAAAATAAGCGGAGCACCTAGACAAAAA	10				
RESULT 10							
CA431219/c							
LOCUS							
DEFINITION							
UI-H-FGI-bgi-i-18-0-UI.s1 NCI CGAP_FGI Homo sapiens cDNA clone							
UI-H-FGI-bgi-i-18-0-UI 3', mRNA sequence.							
CA431219							
VERSION							
CA431219.1 GI:24793945							
KEYWORDS							
EST.							
SOURCE							
Homo sapiens (human)							
ORGANISM							
Homo sapiens							
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;							
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;							
Hominidae; Homo.							
1 (bases 1 to 753)							
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.							
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),							
TITLE							
CA431219/c							
LOCUS							
DEFINITION							
UI-H-FGI-bgi-i-18-0-UI.s1 NCI CGAP_FGI Homo sapiens cDNA clone							
UI-H-FGI-bgi-i-18-0-UI 3', mRNA sequence.							
CA431219							
VERSION							
CA431219.1 GI:24793945							
KEYWORDS							
EST.							
SOURCE							
Homo sapiens (human)							
ORGANISM							
Homo sapiens							
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;							
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;							
Hominidae; Homo.							
1 (bases 1 to 753)							
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.							
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),							
TITLE							
CA431219/c							
LOCUS							
DEFINITION							
UI-H-FGI-bgi-i-18-0-UI.s1 NCI CGAP_FGI Homo sapiens cDNA clone							
UI-H-FGI-bgi-i-18-0-UI 3', mRNA sequence.							
CA431219							
VERSION							
CA431219.1 GI:24793945							
KEYWORDS							
EST.							
SOURCE							
Homo sapiens (human)							
ORGANISM							
Homo sapiens							
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;							
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;							
Hominidae; Homo.							
1 (bases 1 to 753)							
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.							
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),							
TITLE							
CA431219/c							
LOCUS							
DEFINITION							
UI-H-FGI-bgi-i-18-0-UI.s1 NCI CGAP_FGI Homo sapiens cDNA clone							
UI-H-FGI-bgi-i-18-0-UI 3', mRNA sequence.							
CA431219							
VERSION							
CA431219.1 GI:24793945							
KEYWORDS							
EST.							
SOURCE							
Homo sapiens (human)							
ORGANISM							
Homo sapiens							
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;							
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;							
Hominidae; Homo.							
1 (bases 1 to 753)							
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.							
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),							
TITLE							
CA431219/c							
LOCUS							
DEFINITION							
UI-H-FGI-bgi-i-18-0-UI.s1 NCI CGAP_FGI Homo sapiens cDNA clone							
UI-H-FGI-bgi-i-18-0-UI 3', mRNA sequence.							
CA431219							
VERSION							
CA431219.1 GI:24793945							
KEYWORDS							
EST.							
SOURCE							
Homo sapiens (human)							
ORGANISM							
Homo sapiens							
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;							
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;							
Hominidae; Homo.							
1 (bases 1 to 753)							
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.							
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),							
TITLE							
CA431219/c							
LOCUS							
DEFINITION							
UI-H-FGI-bgi-i-18-0-UI.s1 NCI CGAP_FGI Homo sapiens cDNA clone							
UI-H-FGI-bgi-i-18-0-UI 3', mRNA sequence.							
CA431219							
VERSION							
CA431219.1 GI:24793945							
KEYWORDS							
EST.							
SOURCE							
Homo sapiens (human)							
ORGANISM							
Homo sapiens							
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;							
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;							
Hominidae; Homo.							
1 (bases 1 to 753)							
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.							
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),							
TITLE							
CA431219/c							
LOCUS							
DEFINITION							
UI-H-FGI-bgi-i-18-0-UI.s1 NCI CGAP_FGI Homo sapiens cDNA clone							
UI-H-FGI-bgi-i-18-0-UI 3', mRNA sequence.							
CA431219							
VERSION							
CA431219.1 GI:24793945							
KEYWORDS							
EST.							
SOURCE							
Homo sapiens (human)							
ORGANISM							
Homo sapiens							
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;							
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;							
Hominidae; Homo.							
1 (bases 1 to 753)							
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.							
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),							
TITLE							
CA431219/c							
LOCUS							
DEFINITION							
UI-H-FGI-bgi-i-18-0-UI.s1 NCI CGAP_FGI Homo sapiens cDNA clone							
UI-H-FGI-bgi-i-18-0-UI 3', mRNA sequence.							
CA431219							



AA156259.1	GI:1727877	DEFINITION	CR736990 Soares fetal liver spleen lNFLS Homo sapiens cDNA clone
EST.		ACCESSION	IMAGp998E2480 ; IMAGE:108671 5', mRNA sequence.
Homo sapiens		VERSION	CR736990
Homo sapiens		KEYWORDS	EST.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		SOURCE	CR736990.1 GI:51586561
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;		ORGANISM	Homo sapiens (human)
Hominidae; Homo.			Homo sapiens
1 (bases 1 to 428)		REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B.,		AUTHORS	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Chissole, S., Dietrich, N., DuBuque, T., Favello, A., Gish, W.,			Hominidae; Homo.
Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N.,			1 (bases 1 to 543)
Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L.,			Ebert, L., Heil, O., Hennig, S., Korn, B., Neubert, P., Partsch, E.,
Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J.,			Peters, M., Radelof, U. and Schneider, D.
Trevisan, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R.		TITLE	I.M.A.G.E. cDNA Clone Collection
and Marra, M.		JOURNAL	Unpublished (2004)
Generation and analysis of 280,000 human expressed sequence tags		COMMENT	Contact: Inge Arlart
Genome Res. 6 (9), 807-828 (1996)			RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
8889549			Heubnerweg 6, D-14059 Berlin, Germany
Contact: Wilson RK			Email: www.rzpd.de
Washington University School of Medicine			RZPD; IMAGp998E2480.
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108			RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
Tel: 314 286 1800			Contact: Inge Arlart
Fax: 314 286 1810			RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Email: est@watson.wustl.edu			Heubnerweg 6, D-14059 Berlin, Germany
This clone is available royalty-free through LINL ; contact the			Tel: +49 30 32639 100
IMAGE Consortium (info@image.llnl.gov) for further information.			Fax: +49 30 32639 111
Seq primer: -28M13 rev2 from Amersham.			www.rzpd.de
		FEATURES	This clone is available royalty-free from RZPD;
		Location/Qualifiers	contact RZPD (clone@rzpd.de) for further information. Seq primer:
		1..428	M13r, Primer sequence: TTTTCACAGGAACACAGTATGAC.
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="GDB:4621481"	
		/db_xref="taxon:9606"	
		/clone="IMAGE:590095"	
		/dev_stage="umbilical vein, 1 passage"	
		/lab_host="SOLR (kanamycin resistant)"	
		/clone_lib="Stratagene endothelial cell 937223"	
		/note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:	
		XhoI; Cloned unidirectionally. Primer: Oligo dt.	
		Umbilical vein endothelial cells, passaged once. Average	
		insert size: 1.0 kb; Uni-ZAP XR Vector; ~5' adaptor	
		sequence: 5' GAATTCGGCAGGAG 3' ~3' adaptor sequence: 5'	
		CTCGAGTTTTTTTTTTTTTTTTTTT 3'"	
		ORIGIN	
		Query Match	92.0%; Score 270.4; DB 1; Length 428;
		Best Local Similarity	99.6%; Pred. No. 5e-73;
		Matches	271; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
		QY	21 TCCCATGATGCTCTGTCAAGAGTGGACAATCACCTACTGAGGAGCCCAACCCCGCCT 80
		DB	272 TCCCATGATGCTCTGTCAAGAGTGGACAATCACCTACTGAGGAGCCCAACCCCGCCT 213
		QY	81 CTGGATGACCGAATCAGGAGTCTAGAGTTTAACTGCAGCGGAACATGTCATTTCTCTATT 140
		DB	212 CTGGATGACCGAATCAGGAGTCTAGAGTTTAACTGCAGCGGAACATGTCATTTCTCTATT 153
		QY	141 TCTGTGACGCGCCCGCATGCCCGCCACCCACCAAGAGCGCAGAGGGCCGATCATGAC 200
		DB	152 TCTGTGACGCGCCCGCATGCCCGCCACCCACCAAGAGCGCAGAGGGCCGATCATGAC 93
		QY	201 CCTCGTCTCTGCAGGGTGTCTGTACACGTTCTTCTGAAAGCTTTTAGACAGTAACAGTTTG 260
		DB	92 CCTCGTCTCTGCAGGGTGTCTGTACACGTTCTTCTGAAAGCTTTTAGACAGTAACAGTTTG 33
		QY	261 CACATGAAAAATAAAGCGAGCACCTAGACAAA 292
		DB	32 CACATGAAAAATAAAGCGAGCACCTAGACAAA 1
RESULT 12			
CR736990			
LOCUS			



Job time : 3358.89 secs

```
RESULT 15
CB114667
LOCUS      CB114667                650 bp    mRNA    linear    EST 28-JAN-2003
DEFINITION K-EST0158371 L10choCK0 Homo sapiens cDNA clone L10choCK0-1-C03 5',
            mRNA sequence.
ACCESSION  CB114667
VERSION    CB114667.1  GI:27940474
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE  1 (bases 1 to 650)
AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
TITLE      21C Frontier Korean EST Project 2001
JOURNAL    Unpublished (2002)
COMMENT    Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 1 row; C column: 03
            High quality sequence stop: 650.
FEATURES   Location/Qualifiers
            source          1..650
                        /organism="Homo sapiens"
                        /mol_type="mRNA"
                        /db_xref="taxon:9606"
                        /clone="L10choCK0-1-C03"
                        /sex="M"
                        /cell_line="Cho-CK"
                        /lab_host="Top10P"
                        /clone_lib="L10choCK0"
                        /note="Organ: Liver; Vector: pT7T3-Pac; Site_1: EcoRI;
                        Site_2: NotI; The library was contributed by the Soares
                        laboratory and it was constructed as described by Bonaldo,
                        M.F., Lennon, G. and Soares, M.B. (1996), Genome Research
                        6(9): 791-806. RNA was prepared from harvested cell
                        culture."
ORIGIN
Query Match      91.6%; Score 269.4; DB 6; Length 650;
Best Local Similarity 99.6%; Pred. No. 1.1e-72;
Matches 270; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY  21  TCCATGATGCTCTGTCAAGAGTGGACATCACCTACTGAGGAGCCCCACCCCGCCT 80
DB   |||||||
QY  236  TCCCATGATGCTCTGTCAAGAGTGGACATCACCTACTGAGGAGCCCCACCCCGCCT 295
DB   |||||||
QY  81  CTGGATGACCGCAATCAGGACTAGAGTTTAATCGCAGCGGAACATGTCATTCTCTATT 140
DB   |||||||
QY  296  CTGGATGACCGCAATCAGGACTAGAGTTTAATCGCAGCGGAACATGTCATTCTCTATT 355
DB   |||||||
QY  141  TCTGTGACGCGCCCCCATGCCCCACCCACCAAGAGGCGAGGCGCCAGTCATGAC 200
DB   |||||||
QY  356  TCTGTGACGCGCCCCCATGCCCCACCCACCAAGAGGCGAGGCGCCAGTCATGAC 415
DB   |||||||
QY  201  CCTCGTCTCTGAGGGTGTCTGTACACGTTCTTCTGAAAGCTTTAGACAGTAACAGTTTG 260
DB   |||||||
QY  416  CCTCGTCTCTGAGGGTGTCTGTACACGTTCTTCTGAAAGCTTTAGACAGTAACAGTTTG 475
DB   |||||||
QY  261  CACATGAAAAATAAAGCGGACCTAGACAA 291
DB   |||||||
QY  476  CACATGAAAAATAAAGCGGACCTAGACAA 506
DB   |||||||
```

Search completed: December 7, 2005, 09:55:57

*This Page Blank (uspto)*

Result No.	Score	Query			DB	ID	Description
		Match	Length				
1	101.4	98.4	427	6	BD223503	94 human	
2	101.4	98.4	427	6	AR243137	Sequence	
3	101.4	98.4	427	6	AR404318	Sequence	
4	101.4	98.4	427	6	AR576159	Sequence	
5	101.4	98.4	1874	6	CO769657	Sequence	
6	101.4	98.4	2031	6	BD016437	Gene dexi	
7	101.4	98.4	2031	8	AP135600	Homo sapi	
8	101.4	98.4	2034	8	BC037974	Homo sapi	
9	101.4	98.4	2057	6	AC288046	Sequence	
10	100.4	97.5	1277	6	BD223416	94 human	
11	100.4	97.5	1277	6	AR243050	Sequence	
12	100.4	97.5	1277	6	AR404231	Sequence	
13	100.4	97.5	1277	6	AR576072	Sequence	
14	98.4	95.5	136328	8	AC004126	Human Chr	
15	98.4	95.5	177073	14	AP000934	Homo sapi	
16	98.4	95.5	195616	14	AC090344	Homo sapi	
17	98.4	95.5	197156	14	AC090345	Homo sapi	
18	98.4	95.5	197856	8	AP000777	Homo sapi	

FT source 1. .427  
/organism="Homo sapiens (human)"  
LOCATION/Qualifiers  
1. .427  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 98.4%; Score 101.4; DB 6; Length 427;  
Best Local Similarity 99.0%; Pred. No. 4e-23;  
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CGGACAGCGGAGCAGTCCAGCTCTGGTTCTCTCGGTTTATTCTGTAGATGAA 60  
DB 281 CGGACAGCGGAGCAGTCCAGCTCTGGTTCTCTCGGTTTATTCTGTAGATGAA 340  
QY 61 ATGGTTCCCATTAATAAGGGGCATGAGCCCTTCTCACAAAAA 103  
DB 341 ATGGTTCCCATTAATAAGGGGCATGAGCCCTTCTCACAAAAA 383

RESULT 2  
AR243137  
LOCUS AR243137 427 bp DNA linear PAT 20-DEC-2002  
DEFINITION Sequence 107 from patent US 6475753.  
ACCESSION AR243137  
VERSION AR243137.1 GI:27289830  
KEYWORDS Unknown.  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 427)  
AUTHORS Ruben,S.M., Ni,J., Rosen,C.A., Wei,Y.-F., Young,P., Florence,K., Soppet,D.R., Brewer,L.A., Endress,G.A., Carter,K.C., Mucenski,M., Ebner,R., Lafleur,D.W., Olsen,H., Shi,Y., Moore,P.A. and Komatsoulis,G.  
TITLE 94 Human Secreted Proteins  
JOURNAL Patent: US 6475753-A 107 05-NOV-2002;  
Human Genome Sciences, Inc.; Rockville, MD  
FEATURES  
source 1. .427  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Query Match 98.4%; Score 101.4; DB 6; Length 427;  
Best Local Similarity 99.0%; Pred. No. 4e-23;  
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CGGACAGCGGAGCAGTCCAGCTCTGGTTCTCTCGGTTTATTCTGTAGATGAA 60  
DB 281 CGGACAGCGGAGCAGTCCAGCTCTGGTTCTCTCGGTTTATTCTGTAGATGAA 340  
QY 61 ATGGTTCCCATTAATAAGGGGCATGAGCCCTTCTCACAAAAA 103  
DB 341 ATGGTTCCCATTAATAAGGGGCATGAGCCCTTCTCACAAAAA 383

RESULT 3  
AR404318  
LOCUS AR404318 427 bp DNA linear PAT 18-DEC-2003  
DEFINITION Sequence 107 from patent US 6627741.  
ACCESSION AR404318  
VERSION AR404318.1 GI:40152416  
KEYWORDS Unknown.  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 427)  
AUTHORS Ruben,S.M., Ni,J., Rosen,C.A., Wei,Y.-F., Young,P., Florence,K., Soppet,D.R., Brewer,L.A., Endress,G.A., Carter,K.C., Mucenski,M., Ebner,R., Lafleur,D.W., Olsen,H., Shi,Y., Moore,P.A. and Komatsoulis,G.

Komatsoulis,G.  
Antibodies to secreted protein HCEJQ69  
Patent: US 6627741-A 107 30-SEP-2003;  
Human Genome Sciences, Inc.; Rockville, MD  
FEATURES  
source 1. .427  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Query Match 98.4%; Score 101.4; DB 6; Length 427;  
Best Local Similarity 99.0%; Pred. No. 4e-23;  
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CGGACAGCGGAGCAGTCCAGCTCTGGTTCTCTCGGTTTATTCTGTAGATGAA 60  
DB 281 CGGACAGCGGAGCAGTCCAGCTCTGGTTCTCTCGGTTTATTCTGTAGATGAA 340  
QY 61 ATGGTTCCCATTAATAAGGGGCATGAGCCCTTCTCACAAAAA 103  
DB 341 ATGGTTCCCATTAATAAGGGGCATGAGCCCTTCTCACAAAAA 383

RESULT 4  
AR576159  
LOCUS AR576159 427 bp DNA linear PAT 14-DEC-2004  
DEFINITION Sequence 107 from patent US 6774216.  
ACCESSION AR576159  
VERSION AR576159.1 GI:56577705  
KEYWORDS Unknown.  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 427)  
AUTHORS Ruben,S.M., Rosen,C.A. and Lafleur,D.W.  
TITLE Antibodies to secreted protein HCEJQ69  
JOURNAL Patent: US 6774216-A 107 10-AUG-2004;  
Human Genome Sciences, Inc.; Rockville, MD  
FEATURES  
source 1. .427  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Query Match 98.4%; Score 101.4; DB 6; Length 427;  
Best Local Similarity 99.0%; Pred. No. 4e-23;  
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CGGACAGCGGAGCAGTCCAGCTCTGGTTCTCTCGGTTTATTCTGTAGATGAA 60  
DB 281 CGGACAGCGGAGCAGTCCAGCTCTGGTTCTCTCGGTTTATTCTGTAGATGAA 340  
QY 61 ATGGTTCCCATTAATAAGGGGCATGAGCCCTTCTCACAAAAA 103  
DB 341 ATGGTTCCCATTAATAAGGGGCATGAGCCCTTCTCACAAAAA 383

RESULT 5  
CQ769657  
LOCUS CQ769657 1874 bp DNA linear PAT 04-MAR-2004  
DEFINITION Sequence 383 from Patent WO2003058021.  
ACCESSION CQ769657  
VERSION CQ769657.1 GI:45114200  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
REFERENCE 1  
AUTHORS Koening-Hoffman,K., Kazinski,M., Schaefer,R. and Kesper,B.  
TITLE Novel apoptosis-inducing dna sequences  
JOURNAL Patent: WO 2003058021-A 383 17-JUL-2003;

```

FEATURES             Xantos Biomedicine AG (DE)
   source             1. .1874
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"

ORIGIN
Query Match          98.4%; Score 101.4; DB 6; Length 1874;
Best Local Similarity 99.0%; Pred. No. 3.5e-23;
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGACAGACGGCAGCAGTCCAGCTCTGGTTTCCTTCCTCGGTTTATTCGTAGATGAA 60
Db 1707 CGGACAGACGGCAGCAGTCCAGCTCTGGTTTCCTTCCTCGGTTTATTCGTAGATGAA 1766

Qy 61 ATGGTTCCCATAAATAGGGGCATGAGCCCTTCCTCACAAAA 103
Db 1767 ATGGTTCCCATAAATAGGGGCATGAGCCCTTCCTCACAAAA 1809

RESULT 6
BD016437             2031 bp      DNA      linear      PAT 27-AUG-2002
LOCUS                Gene derived from human osteoclast.
DEFINITION           BD016437
ACCESSION            BD016437.1 GI:22557613
VERSION              JP 2001231573-A/3.
KEYWORDS             Homo sapiens (human)
SOURCE               Homo sapiens
ORGANISM             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                    Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
                    Hominiidae; Homo.
REFERENCE            1 (bases 1 to 2031)
AUTHORS             Yamane,M., Maeda,T., Tsuruta,Y., Suzuki,R. and Ochi,T.
TITLE               Cloning of a novel gene of peptide transporter from human
JOURNAL             Osteoclast
PUBLISHED           Unpublished
SUBMITTER            Yamane,S., Toyosaki-Maeda,T., Tsuruta,Y., Suzuki,R. and Ochi,T.
DIRECT SUBMISSION    Submitted (19-MAR-1999) Shionogi Institute for Medical Science,
JOURNAL             Shionogi & Co., Ltd., 2-5-1 Mishima, Settu, Osaka 566-0022, Japan
FEATURES             source
   source            1. .2031
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="11"
                     /map="11q12.2"
                     /cell_type="osteoclast"
                     /tissue_type="RA synovial fluid"
                     1. .2031
                     /gene="OCTP"
                     128..1873
                     /gene="OCTP"
                     /note="genomic sequence found within Homo sapiens
                     chromosome 11q12.2 PAC clone pDJ60696 deposited in Genbank
                     Accession Number AC004126"
                     /codon_start=1
                     /product="osteoclast transporter hOCTP"
                     /protein_id="AAQ13565.1"
                     /db_xref="GI:33337871"
                     /translation="MPAPRARERQPRVPERQPLPRGARGRRWRRAAGAAVLLVEML
                     ERAFFGVGTANLVLYNSTNFTNQEQATRAALVFLGASVLLAPVGGWGLADVILGRYR
                     AVALSLLLYLAASGLLPATAPPDGRSSFCGEMPASPLGACPACSPGSSPSYCAPV
                     LYAGLLLLGLAASVRSNLATSRGADQVMDLGRDATRFFENFYWSINLGAVLSLLVVA
                     FIQGNISFLGYSIPVCVGLAFIFLFPFVFIKPPMGQSVMKLLALQNCPPQL
                     WQHSARDRCARVLADERSPPGASPOEDIANFQVILFVWVTLFVFWVFMQPMQ
                     STVYLOGLHLHPNIPFANPANISVALRAQSSYTIPEAWLLLANVVVLLVLPVKDR
                     LIDPLLRCKLLPSALQKMGFMFGFTSVIVAGVLEMERLHYIHNHETVSQQIGEVL
                     YNAAPLSIWWQIPQYLLIGISEIFASIPGLEFAYSEAPRSMQGAIMGIFPCLSGVGS
                     LGSSIVALLSLPGWMLHCPKDFGNINNCMDLYFFLLAGIQAVTALLFWVIAGRYERA
                     SQGPASHSRFRSDRG"
                     2008
                     /gene="OCTP"

polyA_site
ORIGIN
Query Match          98.4%; Score 101.4; DB 8; Length 2031;
Best Local Similarity 99.0%; Pred. No. 3.5e-23;
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGACAGACGGCAGCAGTCCAGCTCTGGTTTCCTTCCTTCGCTTTATTCGTAGATGAA 60
Db 1909 CGGACAGACGGCAGCAGTCCAGCTCTGGTTTCCTTCGCTTTATTCGTAGATGAA 1968

Qy 61 ATGGTTCCCATAAATAGGGGCATGAGCCCTTCCTCACAAAA 103
Db 1969 ATGGTTCCCATAAATAGGGGCATGAGCCCTTCCTCACAAAA 2011

RESULT 8
BC037974
```

```

FEATURES             Xantos Biomedicine AG (DE)
   source             1. .1874
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"

ORIGIN
Query Match          98.4%; Score 101.4; DB 6; Length 1874;
Best Local Similarity 99.0%; Pred. No. 3.5e-23;
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGACAGACGGCAGCAGTCCAGCTCTGGTTTCCTTCCTCGGTTTATTCGTAGATGAA 60
Db 1707 CGGACAGACGGCAGCAGTCCAGCTCTGGTTTCCTTCCTCGGTTTATTCGTAGATGAA 1766

Qy 61 ATGGTTCCCATAAATAGGGGCATGAGCCCTTCCTCACAAAA 103
Db 1767 ATGGTTCCCATAAATAGGGGCATGAGCCCTTCCTCACAAAA 1809

RESULT 6
BD016437             2031 bp      DNA      linear      PAT 27-AUG-2002
LOCUS                Gene derived from human osteoclast.
DEFINITION           BD016437
ACCESSION            BD016437.1 GI:22557613
VERSION              JP 2001231573-A/3.
KEYWORDS             Homo sapiens (human)
SOURCE               Homo sapiens
ORGANISM             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                    Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
                    Hominiidae; Homo.
REFERENCE            1 (bases 1 to 2031)
AUTHORS             Yamane,M., Maeda,T., Tsuruta,Y., Suzuki,R. and Ochi,T.
TITLE               Gene derived from human osteoclast
JOURNAL             Patent: JP 2001231573-A 3 28-AUG-2001;
                    SHIONOGI & CO LTD
COMMENT             OS Homo sapiens (human)
                    PN JP 2001231573-A/3
                    PD 28-AUG-2001
                    PF 21-FEB-2000 JP 2000048886
                    PI MASAHARU YAMANE,TOMOKO MAEDA,YUJI TSURUTA,RUYUJI SUZUKI, PI
                    TAKAHIRO OCHI
                    PC
                    C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
                    10,
                    PC C12P21/02, C12Q1/68, G01N33/15, G01N33/50// (C12N15/09, C12R1:91) ,
                    PC C12N15/00,
                    CC C12N5/00, (C12N15/00, C12R1:91)
                    Key Key Location/Qualifiers
                    FT CDS Location/Qualifiers
                    1. .2031
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"

ORIGIN
Query Match          98.4%; Score 101.4; DB 6; Length 2031;
Best Local Similarity 99.0%; Pred. No. 3.5e-23;
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGACAGACGGCAGCAGTCCAGCTCTGGTTTCCTTCCTTCGCTTTATTCGTAGATGAA 60
Db 1909 CGGACAGACGGCAGCAGTCCAGCTCTGGTTTCCTTCGCTTTATTCGTAGATGAA 1968

Qy 61 ATGGTTCCCATAAATAGGGGCATGAGCCCTTCCTCACAAAA 103
Db 1969 ATGGTTCCCATAAATAGGGGCATGAGCCCTTCCTCACAAAA 2011
```

LOCUS	BC037974	2034 bp	mRNA	linear	PRI 30-JUN-2004
DEFINITION	Homo sapiens solute carrier family 15, member 3, mRNA (cDNA clone MGC:46085 IMAGE:5213632), complete cds.				
ACCESSION	BC037974				
VERSION	BC037974.1	GI:23273712			
KEYWORDS	MGC.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominoidea; Homo.				
REFERENCE	1 (bases 1 to 2034)				
AUTHORS	Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buettow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaby, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalley, D.B., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.				
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)				
PUBMED	12477932				
REFERENCE	2 (bases 1 to 2034)				
AUTHORS	Strausberg, R.				
TITLE	Direct Submission				
JOURNAL	Submitted (20-SEP-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA				
REMARK	NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a>				
COMMENT	Contact: MGC help desk Email: <a href="mailto:gcaps@mail.nih.gov">gcaps@mail.nih.gov</a> Tissue Procurement: Life Technologies, Inc. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland; Web site: <a href="http://www.nisc.nih.gov/">http://www.nisc.nih.gov/</a> Contact: <a href="mailto:nisc_mgc@nhi.nih.gov">nisc_mgc@nhi.nih.gov</a> Akhter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakeley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaapi, R., Maduro, Q.L., Masiello, C., Mabeiri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.				
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Series: IRAK Plate: 78 Row: 9 Column: 11 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 7706116. Location/Qualifiers 1. .2034 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="MGC:46085 IMAGE:5213632" /tissue_type="Blood, adult leukocytes"				
source	Query Match 98.4%; Score 101.4; DB 6; Length 2057; Best Local Similarity 99.0%; Pred. No. 3.5e-23; Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;				
	QY 1 CGGACAGACGGCAGCAGTCCCAGCTCTGGTTTCTCGGTTTATTTCTGTAGATGAA 60 DB 1914 CGGACAGACGGCAGCAGTCCCAGCTCTGGTTTCTCGGTTTATTTCTGTAGATGAA 1973				



Qy 61 ATGTTTCCATAAATAGGGGCATGAGCCCTTCTCACAATAA 103  
|||||  
Db 1974 ATGTTTCCATAAATAGGGGCATGAGCCCTTCTCAGAAA 2016  
|||||

RESULT 10  
BD223416 1277 bp DNA linear PAT 17-JUL-2003  
LOCUS 94 human secretory proteins.  
DEFINITION BD223416  
ACCESSION BD223416  
VERSION 1 GI:33033186  
KEYWORDS JP 2002518010-A/19.  
SOURCE Homo sapiens (human)  
ORGANISM

REFERENCE 1 (bases 1 to 1277)  
AUTHORS Ruben,S.M., Ni,J., Rosen,C.A., Wei,Y.F., Young,P.E., Florence,K.A., Soppet,D.R., Brewer,L.A., Endress,G.A., Carter,K.C., Mucenski,M., Ebner,R., Lafleur,D.W., Olsen,H.S., Shi,Y., Moore,P.A. and Komatsoulis,G.  
TITLE 94 human secretory proteins  
JOURNAL HUMAN GENOME SCIENCES INC  
COMMENT OS -Homo sapiens (human)  
PN JP 2002518010-A/19  
PD 25-JUN-2002  
PF 15-JUN-1999 JP 2000554850  
PR 16-JUN-1998 US 60/089509,16-JUN-1998 US 60/089510 PR  
16-JUN-1998 US 60/089508,16-JUN-1998 US 60/089507 PR  
22-JUN-1998 US 60/090112,22-JUN-1998 US 60/090113 PI STEVEN M RUBEN, JIAN NI, CRAIG A ROSEN, YING FEI WEI, PAUL E YOUNG  
PI KIMBERLY A FLORENCE, DANIEL R SOPPET, LAURIE A BREWER, GREGORY A PI KENNETH C CARTER, MICHAEL MUCENSKI, REINHARD EBNER, DAVID W PI LAFLEUR,  
PI HENRIK S OLSEN, YANGGU SHI, PAUL A MOORE, GEORGE KOMATSOU LIS PC C12N15/09, A61K38/00, A61K39/395, A61K48/00, A61P43/00, PC C07K7/00, PC C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/10 PC , C12Q1/68, C12N15/00, PC A61K37/02, C12N5/00 CC n equals a,t,g, or c CC n equals a,t,g, or c  
FH Key Location/Qualifiers  
FT source 1..1277  
FT Location/Qualifiers  
/organism='Homo sapiens (human)'.  
/mol\_type='genomic DNA'  
/db\_xref='taxon:9606'

ORIGIN  
Query Match 97.5%; Score 100.4; DB 6; Length 1277;  
Best Local Similarity 98.1%; Pred. No. 7.8e-23;  
Matches 101; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 CGGACAGCGGCAGTCCAGCTCTGGTTTCCTTCGGTTTATTCGTAGAAATGAA 60  
|||||  
Db 1130 CGGACAGCGGCAGTCCAGCTCTGGTTTCCTTCGGTTTATTCGTAGAAATGAA 1189  
|||||

Qy 61 ATGTTTCCATAAATAGGGGCATGAGCCCTTCTCACAATAA 103  
|||||  
Db 1190 ATGTTTCCATAAATAGGGGCATGAGCCCTTCTCACAATAA 1232  
|||||

RESULT 11  
AR243050 1277 bp DNA linear PAT 20-DEC-2002  
LOCUS 94 human secretory proteins.  
DEFINITION AR243050  
ACCESSION AR243050  
VERSION 1 GI:33033186  
KEYWORDS JP 2002518010-A/19.  
SOURCE Homo sapiens (human)  
ORGANISM

REFERENCE 1 (bases 1 to 1277)  
AUTHORS Ruben,S.M., Ni,J., Rosen,C.A., Wei,Y.F., Young,P.E., Florence,K.A., Soppet,D.R., Brewer,L.A., Endress,G.A., Carter,K.C., Mucenski,M., Ebner,R., Lafleur,D.W., Olsen,H.S., Shi,Y., Moore,P.A. and Komatsoulis,G.  
TITLE 94 human secretory proteins  
JOURNAL HUMAN GENOME SCIENCES INC  
COMMENT OS -Homo sapiens (human)  
PN JP 2002518010-A/19  
PD 25-JUN-2002  
PF 15-JUN-1999 JP 2000554850  
PR 16-JUN-1998 US 60/089509,16-JUN-1998 US 60/089510 PR  
16-JUN-1998 US 60/089508,16-JUN-1998 US 60/089507 PR  
22-JUN-1998 US 60/090112,22-JUN-1998 US 60/090113 PI STEVEN M RUBEN, JIAN NI, CRAIG A ROSEN, YING FEI WEI, PAUL E YOUNG  
PI KIMBERLY A FLORENCE, DANIEL R SOPPET, LAURIE A BREWER, GREGORY A PI KENNETH C CARTER, MICHAEL MUCENSKI, REINHARD EBNER, DAVID W PI LAFLEUR,  
PI HENRIK S OLSEN, YANGGU SHI, PAUL A MOORE, GEORGE KOMATSOU LIS PC C12N15/09, A61K38/00, A61K39/395, A61K48/00, A61P43/00, PC C07K7/00, PC C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/10 PC , C12Q1/68, C12N15/00, PC A61K37/02, C12N5/00 CC n equals a,t,g, or c CC n equals a,t,g, or c  
FH Key Location/Qualifiers  
FT source 1..1277  
FT Location/Qualifiers  
/organism='Homo sapiens (human)'.  
/mol\_type='genomic DNA'  
/db\_xref='taxon:9606'

ORIGIN  
Query Match 97.5%; Score 100.4; DB 6; Length 1277;  
Best Local Similarity 98.1%; Pred. No. 7.8e-23;  
Matches 101; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 CGGACAGCGGCAGTCCAGCTCTGGTTTCCTTCGGTTTATTCGTAGAAATGAA 60  
|||||  
Db 1130 CGGACAGCGGCAGTCCAGCTCTGGTTTCCTTCGGTTTATTCGTAGAAATGAA 1189  
|||||

Qy 61 ATGTTTCCATAAATAGGGGCATGAGCCCTTCTCACAATAA 103  
|||||  
Db 1190 ATGTTTCCATAAATAGGGGCATGAGCCCTTCTCACAATAA 1232  
|||||

RESULT 13  
AR243050 1277 bp DNA linear PAT 18-DEC-2003  
LOCUS 94 human secretory proteins.  
DEFINITION AR243050  
ACCESSION AR243050  
VERSION 1 GI:33033186  
KEYWORDS JP 2002518010-A/19.  
SOURCE Homo sapiens (human)  
ORGANISM

REFERENCE 1 (bases 1 to 1277)  
AUTHORS Ruben,S.M., Ni,J., Rosen,C.A., Wei,Y.F., Young,P.E., Florence,K.A., Soppet,D.R., Brewer,L.A., Endress,G.A., Carter,K.C., Mucenski,M., Ebner,R., Lafleur,D.W., Olsen,H.S., Shi,Y., Moore,P.A. and Komatsoulis,G.  
TITLE 94 human secretory proteins  
JOURNAL HUMAN GENOME SCIENCES INC  
COMMENT OS -Homo sapiens (human)  
PN JP 2002518010-A/19  
PD 25-JUN-2002  
PF 15-JUN-1999 JP 2000554850  
PR 16-JUN-1998 US 60/089509,16-JUN-1998 US 60/089510 PR  
16-JUN-1998 US 60/089508,16-JUN-1998 US 60/089507 PR  
22-JUN-1998 US 60/090112,22-JUN-1998 US 60/090113 PI STEVEN M RUBEN, JIAN NI, CRAIG A ROSEN, YING FEI WEI, PAUL E YOUNG  
PI KIMBERLY A FLORENCE, DANIEL R SOPPET, LAURIE A BREWER, GREGORY A PI KENNETH C CARTER, MICHAEL MUCENSKI, REINHARD EBNER, DAVID W PI LAFLEUR,  
PI HENRIK S OLSEN, YANGGU SHI, PAUL A MOORE, GEORGE KOMATSOU LIS PC C12N15/09, A61K38/00, A61K39/395, A61K48/00, A61P43/00, PC C07K7/00, PC C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/10 PC , C12Q1/68, C12N15/00, PC A61K37/02, C12N5/00 CC n equals a,t,g, or c CC n equals a,t,g, or c  
FH Key Location/Qualifiers  
FT source 1..1277  
FT Location/Qualifiers  
/organism='Homo sapiens (human)'.  
/mol\_type='genomic DNA'  
/db\_xref='taxon:9606'

ORIGIN  
Query Match 97.5%; Score 100.4; DB 6; Length 1277;  
Best Local Similarity 98.1%; Pred. No. 7.8e-23;  
Matches 101; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 CGGACAGCGGCAGTCCAGCTCTGGTTTCCTTCGGTTTATTCGTAGAAATGAA 60  
|||||  
Db 1130 CGGACAGCGGCAGTCCAGCTCTGGTTTCCTTCGGTTTATTCGTAGAAATGAA 1189  
|||||

Qy 61 ATGTTTCCATAAATAGGGGCATGAGCCCTTCTCACAATAA 103  
|||||  
Db 1190 ATGTTTCCATAAATAGGGGCATGAGCCCTTCTCACAATAA 1232  
|||||

RESULT 13  
AR243050 1277 bp DNA linear PAT 18-DEC-2003  
LOCUS 94 human secretory proteins.  
DEFINITION AR243050  
ACCESSION AR243050  
VERSION 1 GI:33033186  
KEYWORDS JP 2002518010-A/19.  
SOURCE Homo sapiens (human)  
ORGANISM

REFERENCE 1 (bases 1 to 1277)  
AUTHORS Ruben,S.M., Ni,J., Rosen,C.A., Wei,Y.F., Young,P.E., Florence,K.A., Soppet,D.R., Brewer,L.A., Endress,G.A., Carter,K.C., Mucenski,M., Ebner,R., Lafleur,D.W., Olsen,H.S., Shi,Y., Moore,P.A. and Komatsoulis,G.  
TITLE 94 human secretory proteins  
JOURNAL HUMAN GENOME SCIENCES INC  
COMMENT OS -Homo sapiens (human)  
PN JP 2002518010-A/19  
PD 25-JUN-2002  
PF 15-JUN-1999 JP 2000554850  
PR 16-JUN-1998 US 60/089509,16-JUN-1998 US 60/089510 PR  
16-JUN-1998 US 60/089508,16-JUN-1998 US 60/089507 PR  
22-JUN-1998 US 60/090112,22-JUN-1998 US 60/090113 PI STEVEN M RUBEN, JIAN NI, CRAIG A ROSEN, YING FEI WEI, PAUL E YOUNG  
PI KIMBERLY A FLORENCE, DANIEL R SOPPET, LAURIE A BREWER, GREGORY A PI KENNETH C CARTER, MICHAEL MUCENSKI, REINHARD EBNER, DAVID W PI LAFLEUR,  
PI HENRIK S OLSEN, YANGGU SHI, PAUL A MOORE, GEORGE KOMATSOU LIS PC C12N15/09, A61K38/00, A61K39/395, A61K48/00, A61P43/00, PC C07K7/00, PC C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/10 PC , C12Q1/68, C12N15/00, PC A61K37/02, C12N5/00 CC n equals a,t,g, or c CC n equals a,t,g, or c  
FH Key Location/Qualifiers  
FT source 1..1277  
FT Location/Qualifiers  
/organism='Homo sapiens (human)'.  
/mol\_type='genomic DNA'  
/db\_xref='taxon:9606'

ORIGIN  
Query Match 97.5%; Score 100.4; DB 6; Length 1277;  
Best Local Similarity 98.1%; Pred. No. 7.8e-23;  
Matches 101; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 CGGACAGCGGCAGTCCAGCTCTGGTTTCCTTCGGTTTATTCGTAGAAATGAA 60  
|||||  
Db 1130 CGGACAGCGGCAGTCCAGCTCTGGTTTCCTTCGGTTTATTCGTAGAAATGAA 1189  
|||||

Qy 61 ATGTTTCCATAAATAGGGGCATGAGCCCTTCTCACAATAA 103  
|||||  
Db 1190 ATGTTTCCATAAATAGGGGCATGAGCCCTTCTCACAATAA 1232  
|||||

AR243050 1277 bp DNA linear PAT 18-DEC-2003  
LOCUS 94 human secretory proteins.  
DEFINITION AR243050  
ACCESSION AR243050  
VERSION 1 GI:33033186  
KEYWORDS JP 2002518010-A/19.  
SOURCE Homo sapiens (human)  
ORGANISM

REFERENCE 1 (bases 1 to 1277)  
AUTHORS Ruben,S.M., Ni,J., Rosen,C.A., Wei,Y.F., Young,P.E., Florence,K.A., Soppet,D.R., Brewer,L.A., Endress,G.A., Carter,K.C., Mucenski,M., Ebner,R., Lafleur,D.W., Olsen,H.S., Shi,Y., Moore,P.A. and Komatsoulis,G.  
TITLE 94 human secretory proteins  
JOURNAL HUMAN GENOME SCIENCES INC  
COMMENT OS -Homo sapiens (human)  
PN JP 2002518010-A/19  
PD 25-JUN-2002  
PF 15-JUN-1999 JP 2000554850  
PR 16-JUN-1998 US 60/089509,16-JUN-1998 US 60/089510 PR  
16-JUN-1998 US 60/089508,16-JUN-1998 US 60/089507 PR  
22-JUN-1998 US 60/090112,22-JUN-1998 US 60/090113 PI STEVEN M RUBEN, JIAN NI, CRAIG A ROSEN, YING FEI WEI, PAUL E YOUNG  
PI KIMBERLY A FLORENCE, DANIEL R SOPPET, LAURIE A BREWER, GREGORY A PI KENNETH C CARTER, MICHAEL MUCENSKI, REINHARD EBNER, DAVID W PI LAFLEUR,  
PI HENRIK S OLSEN, YANGGU SHI, PAUL A MOORE, GEORGE KOMATSOU LIS PC C12N15/09, A61K38/00, A61K39/395, A61K48/00, A61P43/00, PC C07K7/00, PC C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/10 PC , C12Q1/68, C12N15/00, PC A61K37/02, C12N5/00 CC n equals a,t,g, or c CC n equals a,t,g, or c  
FH Key Location/Qualifiers  
FT source 1..1277  
FT Location/Qualifiers  
/organism='Homo sapiens (human)'.  
/mol\_type='genomic DNA'  
/db\_xref='taxon:9606'

ORIGIN  
Query Match 97.5%; Score 100.4; DB 6; Length 1277;  
Best Local Similarity 98.1%; Pred. No. 7.8e-23;  
Matches 101; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 CGGACAGCGGCAGTCCAGCTCTGGTTTCCTTCGGTTTATTCGTAGAAATGAA 60  
|||||  
Db 1130 CGGACAGCGGCAGTCCAGCTCTGGTTTCCTTCGGTTTATTCGTAGAAATGAA 1189  
|||||

Qy 61 ATGTTTCCATAAATAGGGGCATGAGCCCTTCTCACAATAA 103  
|||||  
Db 1190 ATGTTTCCATAAATAGGGGCATGAGCCCTTCTCACAATAA 1232  
|||||

RESULT 13  
AR243050 1277 bp DNA linear PAT 18-DEC-2003  
LOCUS 94 human secretory proteins.  
DEFINITION AR243050  
ACCESSION AR243050  
VERSION 1 GI:33033186  
KEYWORDS JP 2002518010-A/19.  
SOURCE Homo sapiens (human)  
ORGANISM

REFERENCE 1 (bases 1 to 1277)  
AUTHORS Ruben,S.M., Ni,J., Rosen,C.A., Wei,Y.F., Young,P.E., Florence,K.A., Soppet,D.R., Brewer,L.A., Endress,G.A., Carter,K.C., Mucenski,M., Ebner,R., Lafleur,D.W., Olsen,H.S., Shi,Y., Moore,P.A. and Komatsoulis,G.  
TITLE 94 human secretory proteins  
JOURNAL HUMAN GENOME SCIENCES INC  
COMMENT OS -Homo sapiens (human)  
PN JP 2002518010-A/19  
PD 25-JUN-2002  
PF 15-JUN-1999 JP 2000554850  
PR 16-JUN-1998 US 60/089509,16-JUN-1998 US 60/089510 PR  
16-JUN-1998 US 60/089508,16-JUN-1998 US 60/089507 PR  
22-JUN-1998 US 60/090112,22-JUN-1998 US 60/090113 PI STEVEN M RUBEN, JIAN NI, CRAIG A ROSEN, YING FEI WEI, PAUL E YOUNG  
PI KIMBERLY A FLORENCE, DANIEL R SOPPET, LAURIE A BREWER, GREGORY A PI KENNETH C CARTER, MICHAEL MUCENSKI, REINHARD EBNER, DAVID W PI LAFLEUR,  
PI HENRIK S OLSEN, YANGGU SHI, PAUL A MOORE, GEORGE KOMATSOU LIS PC C12N15/09, A61K38/00, A61K39/395, A61K48/00, A61P43/00, PC C07K7/00, PC C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/10 PC , C12Q1/68, C12N15/00, PC A61K37/02, C12N5/00 CC n equals a,t,g, or c CC n equals a,t,g, or c  
FH Key Location/Qualifiers  
FT source 1..1277  
FT Location/Qualifiers  
/organism='Homo sapiens (human)'.  
/mol\_type='genomic DNA'  
/db\_xref='taxon:9606'

ORIGIN  
Query Match 97.5%; Score 100.4; DB 6; Length 1277;  
Best Local Similarity 98.1%; Pred. No. 7.8e-23;  
Matches 101; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 CGGACAGCGGCAGTCCAGCTCTGGTTTCCTTCGGTTTATTCGTAGAAATGAA 60  
|||||  
Db 1130 CGGACAGCGGCAGTCCAGCTCTGGTTTCCTTCGGTTTATTCGTAGAAATGAA 1189  
|||||

Qy 61 ATGTTTCCATAAATAGGGGCATGAGCCCTTCTCACAATAA 103  
|||||  
Db 1190 ATGTTTCCATAAATAGGGGCATGAGCCCTTCTCACAATAA 1232  
|||||

DEFINITION Sequence 20 from patent US 6774216.  
 ACCESSION AR576072 GI:56577618  
 VERSION AR576072.1  
 KEYWORDS Unknown.  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 1277)  
 AUTHORS Ruben,S.M., Rosen,C.A. and LaFleur,D.W.  
 TITLE Antibodies to secreted protein HCEJQ69  
 JOURNAL Patent: US 6774216-A 20 10-AUG-2004;  
 Human Genome Sciences, Inc.; Rockville, MD

FEATURES  
 source 1..1277  
 Location/Qualifiers  
 /organism="unknown"  
 /mol\_type="genomic DNA"

ORIGIN  
 Query Match 97.5%; Score 100.4; DB 6; Length 1277;  
 Best Local Similarity 98.1%; Pred.No. 7.8e-23;  
 Matches 101; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGACAGCGGAGCAGTCCAGCTCTGGTTTCCTTCGGTTTATTCCTGTAGATGAA 60  
 Db 1130 CGGACAGCGGAGCAGTCCAGCTCTGGTTTCCTTCGGTTTATTCCTGTAGATGAA 1189

Qy 61 ATGTTTCCCAATAATAGGGGCGATGAGCCCTTCCTCAAAAA 103  
 Db 1190 ATGTTTCCCAATAATAGGGGCGATGAGCCCTTCCTCAAAAA 1232

RESULT 14  
 AC004126 136328 bp DNA linear PRI 20-JUN-1998  
 LOCUS Human Chromosome 11q12.2 PAC clone PDJ60696, complete sequence.  
 DEFINITION  
 ACCESSION AC004126  
 VERSION AC004126.1 GI:3242744  
 KEYWORDS HTG.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homidae; Homo.  
 1 (bases 1 to 136328)  
 Evans,G.A., Athanasiou,M., Aguayo,P., Armstrong,D., Basit,M.,  
 Buettner,J., Bumeister,R., Card,P., deSailboat,F., Dunn,J.,  
 English,C., Ethridge,S., Garner,H.R., Gee,V., Gordon,M., Gotway,G.,  
 Grant,O., Hahner,L., Harris,J., Lewis,E., Loo,H., Loo,K.N.,  
 Major,T., McFarland,J., Newton,J., Osborne-Lawrence,S.,  
 Schageman,J., Schilling,P., Schultz,R., Syed,M., Valenzuela,D.,  
 Ward,T. and Wilson,R.  
 HTGS Submission  
 Unpublished  
 2 (bases 1 to 136328)  
 Evans,G.A., Athanasiou,M., Basit,M., Bradbury,P., Brignac,S.,  
 Bumeister,R., Davis,C., English,C., Franklin,T.L., Garner,H.R.,  
 Gee,V., Gordon,M., Gotway,G., Grant,O., Hahner,L., Harris,J.,  
 Hinson,S., Narayanasamy,U., Newton,J., O'Brien,K., Patel,P.,  
 Schageman,J., Schilling,P., Schultz,R., Syed,M., Valenzuela,D.,  
 Ward,T. and Wilson,R.  
 Direct Submission  
 Submitted (06-FEB-1998) Genome Science & Technology Center,  
 University of Texas Southwestern Medical Center, 5323 Harry Hines  
 Blvd, Dallas, TX 75235-8591, USA  
 3 (bases 1 to 136328)  
 Evans,G.A., Athanasiou,M., Aguayo,P., Armstrong,D., Basit,M.,  
 Buettner,J., Bumeister,R., Card,P., deSailboat,F., Dunn,J.,  
 English,C., Ethridge,S., Garner,H.R., Gee,V., Gordon,M., Gotway,G.,  
 Grant,O., Hahner,L., Harris,J., Lewis,E., Loo,H., Loo,K.N.,  
 Major,T., McFarland,J., Newton,J., Osborne-Lawrence,S.,  
 Schageman,J., Schilling,P., Schultz,R., Stinson,S., Syed,M. and Ward,T.  
 Direct Submission  
 Submitted (20-JUN-1998) Genome Science & Technology Center,  
 University of Texas Southwestern Medical Center, 5323 Harry Hines

COMMENT  
 Blvd, Dallas, TX 75235-8591, USA  
 On Jun 20, 1998 this sequence version replaced gi:2842778.  
 IMPORTANT: This submission contains the entire insert of clone  
 PDJ60696. PDJ60696 comes from a PAC library constructed at the  
 Roswell Park Cancer Institute by the Pieter de Jong group. This  
 clone has been finished according to strict quality criteria and  
 attempts have been made to resolve all base calling problems such  
 as compressions and repetitive elements. The expected Phred/Phrap  
 calculated errors/10kb is 0.44. In addition, attempts have been  
 made to assure over 99% of consensus base calls consist of either  
 double-stranded coverage or 2 types of labeling chemistry on one  
 strand.  
 CHROMOSOMAL LOCUS: This PAC clone comes from the Chromosome  
 11p12.2 Best's disease region mapped between STS D11S461 and EST  
 AHNK. This region spans over 1.5 Mbp.  
 MARKER CONFIRMATION: Sequence confirmed STSs; D11S461, SHGC-36412,  
 WI-12296  
 MAPPED CLONE OVERLAP: HTGS phase 1 PAC clone PDJ1081b4.

FEATURES  
 source 1..136328  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 3..111  
 repeat\_region /rpt\_family="Alu"  
 repeat\_region 4399..4574  
 repeat\_region /rpt\_family="Alu"  
 repeat\_region 5447..5700  
 repeat\_region /rpt\_family="Alu"  
 repeat\_region 6418..6708  
 repeat\_region /rpt\_family="Alu"  
 repeat\_region 8866..9147  
 repeat\_region /rpt\_family="Alu"  
 repeat\_region 10302..10604  
 repeat\_region /rpt\_family="Alu"  
 repeat\_region 10827..11119  
 repeat\_region /rpt\_family="Alu"  
 repeat\_region complement(12208..12498)  
 repeat\_region /rpt\_family="Alu"  
 repeat\_region complement(13578..13749)  
 repeat\_region /rpt\_family="MER39"  
 repeat\_region complement(13647..13737)  
 repeat\_region /rpt\_family="MER21"  
 repeat\_region complement(13805..13947)  
 repeat\_region /rpt\_family="MER39"  
 repeat\_region complement(13822..13908)  
 repeat\_region /rpt\_family="MER21"  
 repeat\_region complement(14283..14759)  
 repeat\_region /rpt\_family="MLT1"  
 repeat\_region 14798..14988  
 repeat\_region /rpt\_family="MER20"  
 repeat\_region complement(15469..15753)  
 repeat\_region /rpt\_family="Alu"  
 repeat\_region complement(16650..16934)  
 repeat\_region /rpt\_family="Alu"  
 repeat\_region complement(18976..19291)  
 repeat\_region /rpt\_family="Alu"  
 repeat\_region 19340..19651  
 repeat\_region /rpt\_family="Alu"  
 repeat\_region 20030..20309  
 repeat\_region /rpt\_family="Alu"  
 repeat\_region 20780..21041  
 repeat\_region /rpt\_family="Alu"  
 repeat\_region complement(23675..23976)  
 repeat\_region /rpt\_family="Alu"  
 repeat\_region complement(24041..24140)  
 repeat\_region /rpt\_family="MER"  
 repeat\_region complement(25164..25433)  
 repeat\_region /rpt\_family="Alu"  
 repeat\_region 26572..26622  
 repeat\_region /rpt\_family="MER5"  
 repeat\_region 27020..27310  
 repeat\_region /rpt\_family="Alu"

```
repeat_region 28214..28492
/rpt family="Alu"
complement(30334..30612)
repeat_region 106147..106441
/rpt family="MIR"
complement(97690..97823)
repeat_region 108079..108291
/rpt family="Alu"
complement(97690..97823)
repeat_region 108358..108646
/rpt family="L1"
complement(109058..109358)
repeat_region 108668..109058
/rpt family="Alu"
complement(109810..109942)
repeat_region 108668..109058
/rpt family="L1"
complement(109810..109942)

Query Match 95.5% Score 98.4; DB 8; Length 136328;
Best Local Similarity 99.0%; Pred.No.2.5e-22;
Matches 99; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGACAGACGGCAGCAGCTCCAGCTCTGGTTTCCTTCGCTTTATTTCTGTTAGATGAA 60
Db 37299 CGGACAGACGGCAGCAGCTCCAGCTCTGGTTTCCTTCGCTTTATTTCTGTTAGATGAA 37358

Qy 61 ATGTTTCCCATTAATAAGGGCAGTCCAGCTTCCTTCGCTTTATTTCTGTTAGATGAA 100
Db 37359 ATGTTTCCCATTAATAAGGGCAGTCCAGCTTCCTTCGCTTTATTTCTGTTAGATGAA 37398

RESULT 15
AP000934 177073 bp DNA linear HTG 30-MAY-2000
LOCUS Homo sapiens chromosome 11 clone RP11-804B24 map 11q13, WORKING
DEFINITION DRAFT SEQUENCE, 43 unordered pieces.
ACCESSION AP000934
VERSION AP000934.2 GI:8119075
KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 177073)
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Homo sapiens 177,073 genomic DNA of 11q13
Published Only in Database (1999)
TITLE Homo sapiens 177,073 genomic DNA of 11q13
JOURNAL Published Only in Database (1999)
REFERENCE 2 (bases 1 to 177073)
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (17-DEC-1999) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
Kitasato Univ., 1-15-1 Kitasato, Sagaminara, Kanagawa 228-8555,
Japan (E-mail:hattori@gsc.riken.go.jp,
URL:http://hgp.gsc.riken.go.jp/, Tel:81-42-778-9923,
Fax:81-42-778-9924)
On May 31, 2000 this sequence version replaced gi:6997748.
COMMENT ----- Genomic Sciences Center(GSC)
Center code: RIKEN
Web site: http://hgp.gsc.riken.go.jp/
Contact: hattori@gsc.riken.go.jp
----- Project Information
Center project name: Humdraft11
Center clone name: RP11-804B24
----- Summary Statistics
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 151872 bases at least Q40
Consensus quality: 164006 bases at least Q30
Consensus quality: 169903 bases at least Q20
Insert size: 172873; sum-of-contigs
```

-----  
Quality coverage: 4.18x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 43 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved

1 18996 contig of 18996 bp in length  
19097 30931 contig of 11835 bp in length  
31032 41792 contig of 10761 bp in length  
41893 52059 contig of 10167 bp in length  
52160 61534 contig of 9375 bp in length  
61635 66662 contig of 5028 bp in length  
66763 74887 contig of 8125 bp in length  
74988 81286 contig of 6299 bp in length  
81387 87963 contig of 6577 bp in length  
88064 92858 contig of 4795 bp in length  
92959 98036 contig of 5078 bp in length  
98137 102219 contig of 4083 bp in length  
102320 107130 contig of 4811 bp in length  
107231 111269 contig of 4039 bp in length  
111370 114943 contig of 3574 bp in length  
115044 120076 contig of 5033 bp in length  
120177 123973 contig of 3797 bp in length  
124074 125880 contig of 3321 bp in length  
125881 129227 contig of 3347 bp in length  
129228 132648 contig of 3321 bp in length  
132649 135758 contig of 3010 bp in length  
135759 138433 contig of 2575 bp in length  
138434 140537 contig of 2004 bp in length  
140538 142449 contig of 1812 bp in length  
142450 145260 contig of 2711 bp in length  
145261 148364 contig of 3574 bp in length  
148365 151360 contig of 2326 bp in length  
151361 154160 contig of 100 bp in length  
154161 157967 contig of 1845 bp in length  
157968 160883 contig of 1239 bp in length  
160884 163065 contig of 1525 bp in length  
163066 166494 contig of 1529 bp in length  
166495 169798 contig of 100 bp in length  
169799 172522 contig of 1627 bp in length  
172523 174311 contig of 1689 bp in length  
174312 175674 contig of 1263 bp in length  
175675 177073 contig of 1299 bp in length

Sequence updated (26-May-2000).

NOTE: This is a 'working draft' sequence. It currently consists of 43 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 18996: contig of 18996 bp in length  
18997 19096: gap of 100 bp  
19097 30931: contig of 11835 bp in length  
30932 31031: gap of 100 bp  
31032 41792: contig of 10761 bp in length  
41793 41892: gap of 100 bp  
41893 52059: contig of 10167 bp in length  
52060 52159: gap of 100 bp  
52160 61534: contig of 9375 bp in length  
61535 61635: gap of 100 bp  
61635 66662: contig of 5028 bp in length  
66663 66762: gap of 100 bp

66763 74887: contig of 8125 bp in length  
74888 74987: gap of 100 bp  
74988 81286: contig of 6299 bp in length  
81287 81386: gap of 100 bp  
81387 87963: contig of 6577 bp in length  
87964 88063: gap of 100 bp  
88064 92858: contig of 4795 bp in length  
92859 98036: contig of 5078 bp in length  
98037 98136: gap of 100 bp  
98137 102219: contig of 4083 bp in length  
102220 102319: gap of 100 bp  
102320 107130: contig of 4811 bp in length  
107131 107230: gap of 100 bp  
107231 111269: contig of 4039 bp in length  
111270 111369: gap of 100 bp  
111370 114943: contig of 3574 bp in length  
114944 115044: gap of 100 bp  
115044 120076: contig of 5033 bp in length  
120077 120176: gap of 100 bp  
120177 123973: contig of 3797 bp in length  
123974 124073: gap of 100 bp  
124074 125880: contig of 1707 bp in length  
125881 129227: contig of 3347 bp in length  
129228 132648: contig of 3321 bp in length  
132649 135758: contig of 3010 bp in length  
135759 138433: contig of 2575 bp in length  
138434 140537: contig of 2004 bp in length  
140538 142449: contig of 1812 bp in length  
142450 145260: contig of 2711 bp in length  
145261 148364: contig of 3574 bp in length  
148365 151360: contig of 2326 bp in length  
151361 154160: gap of 100 bp  
154161 157967: contig of 1845 bp in length  
157968 160883: contig of 1239 bp in length  
160884 163065: contig of 1525 bp in length  
163066 166494: contig of 1529 bp in length  
166495 169798: contig of 100 bp in length  
169799 172522: contig of 1627 bp in length  
172523 174311: contig of 1689 bp in length  
174312 175674: contig of 1263 bp in length  
175675 177073: contig of 1299 bp in length

```
FEATURES
  source      Location/Qualifiers
1..177073    /organism="Homo sapiens"
             /mol_type="genomic DNA"
             /db_xref="taxon:9606"
             /chromosome="11"
             /map="11q13"
             /clone="RP11-804B24"
misc_feature 1..18996    /note="assembly_fragment"
misc_feature 19097..30931

Query Match      95.5%; Score 98.4; DB 14; Length 177073;
Best Local Similarity 99.0%; Pred. No. 2.4e-22;
Matches 99; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 CGGACAGCGGACGAGTCCGAGCTCTGGTTTCCTTCTCGGTTTATTCTGTAGATGAA 60
        |||
Db      30399 CGGACAGCGGACGAGTCCGAGCTCTGGTTTCCTTCTCGGTTTATTCTGTAGATGAA 30458
        |||

Qy      61 ATGGTTCCCATATAAGGGGCGATGAGCCCTTCTCACA 100
        |||
Db      30459 ATGGTTCCCATATAAGGGGCGATGAGCCCTTCTCACA 30498
        |||
```

Search completed: December 7, 2005, 04:22:11  
Job time : 913.919 secs

**This Page Blank (uspio)**

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 11:47:47 ; Search time 161.924 Seconds  
(without alignments) 4239.411 Million cell updates/sec

**Title:** US-09-980-046B-8

Perfect score: 103

Perfect score: 100  
Sequence: 1 cagacagacagcagcaqtcc.....tgaqcccttcctcacaaaa 103

Scoring table: IDENTITY NUC

scoring table: IDENTITY\_NOC  
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%

FOOT-PROCESSING: Minimum Match 0%  
Maximum Match 100%

### Listing first 45 summaries

Database : N Geneseq 21:\*

11: geneseqn1980as:\*

12: geneseqn1980as:\*

13: geneseqn1998as:\*

14: geneseqn2008as:\*

15: geneseqn2001as:

16: geneseqn2001bs:

17: geneseqn2002as:

18: geneseqn2002bs:

19: geneseqn2003as:

20: geneseqn2003bs:

21: geneseqn2003cs:

22: geneseqn2003ds:

23: geneseqn2004as:

24: geneseqn2004bs:

25: geneseqn2005as:

26: geneseqn2005bs:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	103	100.0	103	4	AAC89739	AAC89739	Human gas
2	102.6	99.6	935	6	AA034122	AA034122	Human sec
3	102.2	99.2	1147	6	AA034143	AA034143	Human sec
4	102.2	99.2	1149	6	AB189938	AB189938	Human pol
5	101.4	98.4	427	3	AA297115	AA297115	Human sec
6	101.4	98.4	427	8	AD403310	AD403310	Human sec
7	101.4	98.4	427	9	ACH66744	ACH66744	Novel hum
8	101.4	98.4	427	9	AD891346	AD891346	Human sec
9	101.4	98.4	439	6	AA034144	AA034144	Human sec
c 10	101.4	98.4	1148	4	AAK56586	AAK56586	Human imm
11	101.4	98.4	1874	10	AD162940	AD162940	Human apo
12	101.4	98.4	2031	5	AA168202	AA168202	Human ost
13	101.4	98.4	2057	13	ADQ88225	ADQ88225	Human 335
14	101.4	98.4	2060	6	AA170679	AA170679	Human mem
15	101.4	98.4	2124	3	AAK52347	AAK52347	Human pol
16	100.4	97.5	1277	3	AAZ97028	AAZ97028	Human sec
17	100.4	97.5	1277	8	AD439759	AD439759	Human sec
18	100.4	97.5	1277	9	ACH66657	ACH66657	Novel hum
19	100.4	97.5	1277	9	AD891137	AD891137	Human sec

## ALIGNMENTS

RESULT 1	
AAC89739	AAC89739 standard; cDNA; 103 BP.
AC	XX
AC	AAC89739;
XX	
XX	12-MAR-2001 (first entry)
XX	
XX	Human gastrointestinal inflammation-related cDNA, SEQ ID NO: 8.
DE	XX
XX	Human; cytostatic; immunomodulator; immunostimulant; vulnery;
KW	anti-inflammatory; neuroprotective; antibacterial; gene therapy;
KW	gastrointestinal inflammation; immune system disorder; genetic disorder;
KW	cancer; autoimmune disorder; infection; wound healing; ss.
XX	
OS	Homo sapiens.
XX	
XX	WO200073324-A2.
FN	
XX	07-DEC-2000.
PD	
XX	
XX	01-JUN-2000; 2000WO-US015191.
PF	
XX	
XX	01-JUN-1999; 99US-0137058P.
PR	
XX	
XX	(DIGI-) DIGITAL GENE TECHNOLOGIES INC.
PA	
XX	
PI	Youakim A, Dubose RF, Sims JE, Pribyl TW, Hillbush BS, Hasel KW;
XX	
XX	WPI; 2001-061508/07.
XX	
XX	New polynucleotides and polypeptides, useful in gene therapy and in
PT	diagnosing a pathological condition, e.g. for modulating gene expression
PT	in gastrointestinal inflammation, or for treating cancers or genetic
PT	disorders.
XX	
XX	Claim 1; Page 85; 108pp; English.
PS	
XX	
CC	The present sequence is one of a number of isolated human polynucleotides
CC	which are useful in gene therapy, and for diagnosing a pathological
CC	condition or a susceptibility to it. In particular, the polynucleotides
CC	are useful for modulating gene expression in gastrointestinal

CC inflammation. The polynucleotides are useful for chromosome  
 CC identification, controlling gene expression through triple helix  
 CC formation or antisense DNA or RNA, or identifying individuals from minute  
 CC biological samples using DNA-based identification techniques. The  
 CC polynucleotides can also be used as an alternative to restriction  
 CC fragment length polymorphism (RFLP), by determining the actual base-by-  
 CC base DNA sequences of selected portions of an individual's genome. The  
 CC polynucleotides may also be used as molecular weight markers on Southern  
 CC gels, as diagnostic probes for the presence of a specific mRNA, as a  
 CC probe to substract-out known sequences in the process of discovering novel  
 CC polynucleotides, or as an antigen to elicit an immune response. The  
 CC polypeptides are useful in diagnostic procedures to detect a disorder.  
 CC The polynucleotides and polypeptides are useful for preventing, treating  
 CC or ameliorating immune system disorders, genetic disorders, cancers, some  
 CC autoimmune disorders, or infections. The polynucleotides and polypeptides  
 CC are also useful for differentiating, proliferating or attracting cells,  
 CC leading to the regeneration of tissues, especially in wounds or burns.  
 CC The polypeptides and polynucleotides may also be used as a food additive  
 CC or preservative  
 XX  
 SQ Sequence 103 BP; 27 A; 26 C; 23 G; 27 T; 0 U; 0 Other;

Query Match 100.0%; Score 103; DB 4; Length 103;  
 Best Local Similarity 100.0%; Pred. No. 4.7e-28;  
 Matches 103; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CGGACAGCGGAGCAGCTCCAGCTCTGGTTCTCTCGGTTATTCGTAGATGAA 60  
 DB 1 CGGACAGCGGAGCAGCTCCAGCTCTGGTTCTCTCGGTTATTCGTAGATGAA 60  
 QY 61 ATGTTTCCCATTAATAAGGGCGCATGAGCCCTTCTCACAAAAA 103  
 DB 61 ATGTTTCCCATTAATAAGGGCGCATGAGCCCTTCTCACAAAAA 103

RESULT 2  
 AAD341122  
 ID AAD341122 standard; cDNA; 935 BP.  
 AC AAD341122;  
 XX 16-JUL-2002 (first entry)  
 XX Human secreted protein-encoding gene 17 cDNA clone HNHOD23, SEQ ID NO:27.  
 XX Human; secreted protein; proliferative disorder; cancer; tumour; AIDS;  
 KW foetal abnormality; developmental; haematopoietic disorder; kidney; skin;  
 KW immune system; cardiovascular; angiogenic; gastrointestinal; endocrine;  
 KW pregnancy-related disorder; cognitive; neurological; autoimmune disease;  
 KW inflammation; allergy; rheumatoid arthritis; Alzheimer's disease; asthma;  
 KW Parkinson's disease; schizophrenia; psoriasis; sepsis; atherosclerosis;  
 KW diabetes; infection; wound healing; vulnery; chemotaxis; food additive;  
 KW gene therapy; gene; ss.

OS Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 FT CDS 466..741  
 FT /\*tag= a  
 FT /product= "Human secreted protein"  
 FT sig\_peptide 466..567  
 FT /\*tag= b  
 FT mat\_peptide 568..738  
 FT /\*tag= c  
 FT /product= "Mature secreted protein"  
 XX  
 XX WO200222654-A1.  
 XX  
 XX 21-MAR-2002.  
 XX  
 XX 17-JAN-2001; 2001WO-US001385.  
 XX  
 XX 12-SEP-2000; 2000US-0231969P.

XX (HUMA-) HUMAN GENOME SCI INC.  
 XX Rosen CA, Komatsoulis GA, Baker KP, Birse CE, Soppet DR;  
 PI Olsen HS, Moore PA, Wei P, Ebner R, Duan DR, Shi Y, Choi GH;  
 PI Fiscella M, Ni J;  
 XX WPI; 2002-315684/35.  
 DR P-PSDB; AAE21594.  
 XX Seventeen nucleic acid molecules encoding human secreted proteins, useful  
 FT for treating and preventing cancer, immune disorders (e.g. Addison's  
 FT disease, and allergies), and cardiovascular disorders (e.g. myocardial  
 FT ischemias).  
 XX Claim 1; Page 423; 483pp; English.  
 XX AAD34106-AAD34144 represent cDNAs corresponding to 17 human secreted  
 CC protein genes, and AAE21578-AAE21616 represent the proteins they encode.  
 CC AAE21617-AAE21627 represent human secreted protein fragments. The genes  
 CC and their corresponding secreted proteins are useful for preventing,  
 CC treating or ameliorating medical conditions, e.g., by protein or gene  
 CC therapy. Pathological conditions can be diagnosed by determining the  
 CC amount of the new protein in a sample or by determining the presence of  
 CC mutations in the new genes. Specific uses are described for each of the  
 CC 17 genes, based on the tissues in which they are most highly expressed,  
 CC and include developing products for the diagnosis or treatment of cancer,  
 CC proliferative disorders, tumours, foetal and developmental abnormalities,  
 CC haematopoietic disorders, diseases of the immune system, AIDS, autoimmune  
 CC diseases (rheumatoid arthritis), inflammation, allergies, neurological  
 CC disorders (Alzheimer's disease, Parkinson's disease), cognitive disorder,  
 CC schizophrenia, asthma, skin disorders (psoriasis), sepsis, diabetes,  
 CC atherosclerosis, cardiovascular disorders, angiogenic disorders, kidney  
 CC disorders, gastrointestinal disorders, pregnancy-related disorders,  
 CC endocrine disorders, and infections. The proteins can also be used to aid  
 CC wound healing and epithelial cell proliferation, to prevent skin aging  
 CC due to sunburn, to maintain organs before transplantation, for supporting  
 CC cell culture of primary tissues, to regenerate tissues, to identify their  
 CC cognate ligands or binding partners, and in chemotaxis, and can be used  
 CC as a food additive or preservative to modify storage properties. The  
 CC present sequence represents a human secreted protein-encoding cDNA of the  
 CC invention  
 XX  
 SQ Sequence 935 BP; 237 A; 261 C; 241 G; 186 T; 0 U; 10 Other;  
 Query Match 99.6%; Score 102.6; DB 6; Length 935;  
 Best Local Similarity 99.0%; Pred. No. 1.6e-27;  
 Matches 102; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CGGACAGCGGAGCAGCTCCAGCTCTGGTTCTCTCGGTTATTCGTAGATGAA 60  
 DB 777 CGGACAGCGGAGCAGCTCCAGCTCTGGTTCTCTCGGTTATTCGTAGATGAA 836  
 QY 61 ATGTTTCCCATTAATAAGGGCGCATGAGCCCTTCTCACAAAAA 103  
 DB 837 ATGTTTCCCATTAATAAGGGCGCATGAGCCCTTCTCACAAAAA 879  
 RESULT 3  
 AAD341143  
 ID AAD341143 standard; cDNA; 1147 BP.  
 XX AAD341143;  
 AC AAD341143;  
 XX 16-JUL-2002 (first entry)  
 XX Human secreted protein-encoding gene 17 cDNA clone HNHOD23, SEQ ID NO:48.  
 XX Human; secreted protein; proliferative disorder; cancer; tumour; AIDS;  
 KW foetal abnormality; developmental; haematopoietic disorder; kidney; skin;  
 KW immune system; cardiovascular; angiogenic; gastrointestinal; endocrine;  
 KW pregnancy-related disorder; cognitive; neurological; autoimmune disease;  
 KW inflammation; allergy; rheumatoid arthritis; Alzheimer's disease; asthma;



KW Parkinson's disease; schizophrenia; psoriasis; sepsis; atherosclerosis;  
KW diabetes; infection; wound healing; vulnery; chemotaxis; food additive;  
KW gene therapy; gene; ss.  
XX Homo sapiens.  
OS  
FH Location/Qualifiers  
FT CDS 678..953  
FT FT /\*tag= a  
FT FT /product= "Human secreted protein"  
FT sig\_peptide 678..779  
FT FT /\*tag= b  
FT mat\_peptide 780..950  
FT FT /\*tag= c  
FT FT /product= "Mature secreted protein"  
XX  
XX WO200222654-A1.  
XX  
XX 21-MAR-2002.  
XX  
XX 17-JAN-2001; 2001WO-US001385.  
XX  
XX 12-SEP-2000; 2000US-0231969P.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Komatoulis GA, Baker KP, Birse CE, Soppet DR;  
PI Olsen HS, Moore PA, Wei P, Ebner R, Duan DR, Shi Y, Choi GH;  
PI Fiscella M, Ni J;  
XX  
XX WPI; 2002-315684/35.  
XX P-PSDB; AAE21615.  
XX  
XX Seventeen nucleic acid molecules encoding human secreted proteins, useful  
FT for treating and preventing cancer, immune disorders (e.g. Addison's  
FT disease, and allergies), and cardiovascular disorders (e.g. myocardial  
FT ischermias).  
XX  
XX Claim 1; Page 436; 483pp; English.  
XX  
XX AAD34106-AAD34144 represent cDNAs corresponding to 17 human secreted  
CC protein genes, and AAE21578-AAE21616 represent the proteins they encode.  
CC AAE21617-AAE21627 represent human secreted protein fragments. The genes  
CC and their corresponding secreted proteins are useful for preventing,  
CC treating or ameliorating medical conditions, e.g., by protein or gene  
CC therapy. Pathological conditions can be diagnosed by determining the  
CC amount of the new protein in a sample or by determining the presence of  
CC mutations in the new genes. Specific uses are described for each of the  
CC 17 genes, based on the tissues in which they are most highly expressed,  
CC and include developing products for the diagnosis or treatment of cancer,  
CC proliferative disorders, tumours, foetal and developmental abnormalities,  
CC haematopoietic disorders, diseases of the immune system, AIDS, autoimmune  
CC diseases (rheumatoid arthritis), inflammation, allergies, neurological  
CC disorders (Alzheimer's disease, Parkinson's disease), cognitive disorder,  
CC schizophrenia, asthma, skin disorders (psoriasis), sepsis, diabetes,  
CC atherosclerosis, cardiovascular disorders, angiogenic disorders, kidney  
CC disorders, gastrointestinal disorders, pregnancy-related disorders,  
CC endocrine disorders, and infections. The proteins can also be used to aid  
CC wound healing and epithelial cell proliferation, to prevent skin aging  
CC due to sunburn, to maintain organs before transplantation, for supporting  
CC cell culture of primary tissues, to regenerate tissues, to identify their  
CC cognate ligands or binding partners, and in chemotaxis, and can be used  
CC as a food additive or preservative to modify storage properties. The  
CC present sequence represents a human secreted protein-encoding cDNA of the  
CC invention  
XX  
XX Sequence 1147 BP; 249 A; 354 C; 286 G; 251 T; 0 U; 7 Other;  
SQ  
Query Match 99.2%; Score 102.2; DB 6; Length 1147;  
Best Local Similarity 98.1%; Pred. No. 2.5e-27;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CGGACAGACGGCAGCAGTCCAGCTCTGGTTTCTTCTCGGTTTATCTGTTAGATGAA 60

Db 989 CGGACAGACGGCAGCAGTCCAGCTCTGGTTTCTTCTCGGTTTATCTGTTAGATGAA 1048  
QY 61 ATGGTTTCCCAATAAATAAGGGGCATGAGCCCTTCTCACAAAA 103  
Db 1049 ATGGTTTCCCAATAAATAAGGGGCATGAGCCCTTCTCAMRAAAA 1091  
RESULT 4  
ABL89938  
ID ABL89938 standard; cDNA; 1149 BP.  
XX  
AC ABL89938;  
XX  
DT 24-MAY-2002 (first entry)  
XX  
DE Human polynucleotide SEQ ID NO 500.  
XX  
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;  
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;  
KW vulnery; anticonvulsant; antibacterial; antifungal; antiparasitic;  
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;  
KW neurological disease; infection; human; secreted protein; gene; ss.  
XX  
XX Homo sapiens.  
XX OS  
XX WO200190304-A2.  
XX  
XX 29-NOV-2001.  
XX  
XX 18-MAY-2001; 2001WO-US016450.  
XX  
XX 19-MAY-2000; 2000US-0205515P.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Birse CE, Rosen CA;  
XX  
XX WPI; 2002-122018/16.  
XX P-PSDB; ABB89529.  
XX  
XX Novel 1405 isolated polypeptides, useful for diagnosis, treatment and  
FT prevention of neural, immune system, muscular, reproductive,  
FT gastrointestinal, pulmonary, cardiovascular, renal and proliferative  
FT disorders.  
XX  
XX Claim 4; SEQ ID NO 500; 2081pp + Sequence Listing; English.  
XX  
XX The invention relates to novel genes (ABL89449-ABL90853) and proteins  
CC (ABB89040-ABB90444) useful for preventing, treating or ameliorating  
CC medical conditions e.g. by protein or gene therapy. The genes are  
CC isolated from a range of human tissues disclosed in the specification.  
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in  
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and  
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,  
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune  
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic  
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,  
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)  
CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing  
CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)  
CC infectious diseases such as viral, bacterial, fungal and parasitic  
CC infections. Note: The sequence data for this patent did not form part of  
CC the printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pat\_sequences  
XX  
XX Sequence 1149 BP; 250 A; 348 C; 292 G; 252 T; 0 U; 7 Other;  
SQ  
Query Match 99.2%; Score 102.2; DB 6; Length 1149;  
Best Local Similarity 98.1%; Pred. No. 2.5e-27;  
Matches 101; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CGGACAGACGGCAGCAGTCCAGCTCTGGTTTCTTCTCGGTTTATCTGTTAGATGAA 60

Db 991 CGGACAGCGCAGCAGCTCCAGCTCTGGTTTCTCTCTGTTATTCTCTTAGAATGAA 1050  
|||  
Qy 61 ATGGTTCCCATTAATAAGGGGCATGAGCCCTTCTCAGAAAA 103  
|||  
Db 1051 ATGGTTCCCATTAATAAGGGGCATGAGCCCTTCTCAGAAAA 1093  
|||  
RESULT 5  
AAZ97115  
ID AAZ97115 standard; cDNA; 427 BP.  
XX AC  
XX AAZ97115;  
XX 19-APR-2000 (first entry)  
XX Human secreted protein gene 10 cDNA clone HDPWU34, SEQ ID NO:107.  
XX  
XX Human; secreted protein; cancer; tumour; developmental abnormality;  
KW foetal deficiency; blood disorder; immune system disorder; inflammation;  
KW autoimmune disease; allergy; Alzheimer's disease; cognitive disorder;  
KW schizophrenia; arthritis; asthma; psoriasis; sepsis; skin disorder;  
KW atherosclerosis; diabetes; cardiovascular disorder; kidney disorder;  
KW digestive disorder; endocrine disorder; infection; AIDS; leukaemia;  
therapy; ds.  
XX  
XX Homo sapiens.  
XX  
XX WO9966041-A1.  
XX  
XX 23-DEC-1999.  
XX  
XX 15-JUN-1999; 99WO-US013418.  
XX  
PR 16-JUN-1998; 98US-0089507P.  
PR 16-JUN-1998; 98US-0089508P.  
PR 16-JUN-1998; 98US-0089509P.  
PR 16-JUN-1998; 98US-0089510P.  
PR 22-JUN-1998; 98US-0090112P.  
PR 22-JUN-1998; 98US-0090113P.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Ruben SM, Ni J, Rosen CA, Wei Y, Young PE, Florence KA;  
PI Soppet DR, Brewer LA, Endress GA, Carter KC, Mucenski M, Ebner R;  
PI Lafleur DW, Olsen HS, Shi Y, Moore PA, Komatsoulis G;  
XX WPI; 2000-106100/09.  
XX P-PSDB; AAY86311.  
XX  
XX New isolated human genes and the secreted polypeptides they encode,  
PT useful for diagnosis and treatment of e.g. cancers, neurological  
PT disorders, immune diseases, inflammation or blood disorders.  
XX  
XX Claim 1; Page 374; 586pp; English.  
XX  
CC AAZ97019 to AAZ97137 represent 94 isolated human secreted protein genes.  
CC AAY86215 to AAY86333 are the secreted proteins encoded by the 94 human  
CC genes. This sequence represents a fragment of one of the human secreted  
CC proteins. The genes and their corresponding secreted polypeptides are  
CC useful for preventing, treating or ameliorating medical conditions, e.g.,  
CC by protein or gene therapy. Also pathological conditions can be diagnosed  
CC by determining the amount of the new polypeptides in a sample or by  
CC determining the presence of mutations in the new genes. Specific uses are  
CC described for each of the 94 genes, based on which tissues they are most  
CC highly expressed in, and include developing products for the diagnosis or  
CC treatment of cancer, tumours, developmental abnormalities and foetal  
CC deficiencies, blood disorders, diseases of the immune system, autoimmune  
CC diseases, inflammation, allergies, Alzheimer's and cognitive disorders,  
CC schizophrenia, arthritis, asthma, psoriasis, sepsis, skin disorders,  
CC atherosclerosis, diabetes, cardiovascular disorders, kidney disorders,  
CC digestive/endocrine disorders, infections and AIDS. The polypeptides are  
CC also useful for identifying their binding partners. The sequences shown

CC in AAY86334 to AAY86585 represent fragments of the secreted proteins  
XX  
SQ Sequence 427 BP; 115 A; 120 C; 102 G; 90 T; 0 U; 0 Other;  
  
Query Match 98.4%; Score 101.4; DB 3; Length 427;  
Best Local Similarity 99.0%; Pred. No. 3.3e-27;  
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 1 CGGACAGCGCAGCAGTCCAGCTCTGGTTTCTCTCTGTTATTCTCTTAGAATGAA 60  
|||  
Db 281 CGGACAGCGCAGCAGTCCAGCTCTGGTTTCTCTCTGTTATTCTCTTAGAATGAA 340  
|||  
Qy 61 ATGGTTCCCATTAATAAGGGGCATGAGCCCTTCTCAGAAAA 103  
|||  
Db 341 ATGGTTCCCATTAATAAGGGGCATGAGCCCTTCTCAGAAAA 383  
|||  
RESULT 6  
ADA40310  
ID ADA40310 standard; cDNA; 427 BP.  
XX  
XX ADA40310;  
XX AC  
XX 20-NOV-2003 (first entry)  
XX DT  
XX Human secreted protein encoding cDNA.  
XX DE  
XX  
XX Human; secreted protein; cancer; hyperproliferative disorder;  
KW rheumatoid arthritis; autoimmune disorder; haematopoietic disorder;  
KW anaemia; allergic reaction; asthma; cardiovascular disorder;  
KW wound healing; cytostatic; immunosuppressive; neoplastic; neuroprotective;  
KW antiviral; anti-allergic; hepatotropic; antidiabetic; anti-inflammatory;  
KW vulnery; cardiant; gene therapy; ss.  
XX  
XX Homo sapiens.  
XX OS  
XX WO2002102993-A2.  
XX PN  
XX 27-DEC-2002.  
XX PD  
XX  
XX 19-MAR-2002; 2002WO-US008123.  
XX PF  
XX  
XX 21-MAR-2001; 2001US-0277340P.  
XX PR  
XX 19-JUL-2001; 2001US-0306171P.  
XX PR  
XX 13-NOV-2001; 2001US-0331287P.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX PA  
XX  
XX Rosen CA, Ruben SM;  
XX PI  
XX WPI; 2003-175238/17.  
XX DR  
XX  
XX New human secreted proteins and nucleic acid molecules, useful for  
PT preparing a diagnostic or pharmaceutical composition for diagnosing,  
PT preventing or treating cancer or other hyperproliferative disorder,  
PT asthma, allergies or AIDS.  
XX  
XX Claim 9; SEQ ID NO 692; 3205pp; English.  
XX PS  
XX The invention relates to novel genes ADA39629-ADA40565 and proteins  
CC ADA40566-ADA41501 for human secreted proteins, useful for preventing,  
CC treating or ameliorating medical conditions e.g. by protein or gene  
CC therapy. The polypeptides, nucleic acid molecules, antibodies or their  
CC fragments, and agonists or antagonists that bind to the polypeptide are  
CC useful for preparing a diagnostic or pharmaceutical composition for  
CC diagnosing or treating cancer or other hyperproliferative disorder. The  
CC polypeptides and nucleic acid molecules are also useful for detecting,  
CC preventing, diagnosing, prognosticating, treating or ameliorating cancer  
CC or other hyperproliferative disorders including neoplasms, autoimmune  
CC disorders (e.g. diabetes, rheumatoid arthritis, systemic lupus  
CC erythematosus, multiple sclerosis, autoimmune thyroiditis or haemolytic  
CC anaemia), haematopoietic or haematological disorders (e.g. anaemia,  
CC thrombocytopenia), allergic reactions including asthma or eczema,

CC inflammatory disorders (e.g. ischaemia-reperfusion injury, inflammatory  
CC bowel disease or Crohn's disease), neurodegenerative disorders (e.g.  
CC Alzheimer's disease or Parkinson's disease), cardiovascular disorders  
CC (e.g. atherosclerosis, myocarditis), infectious diseases (bacterial,  
CC fungal or viral infections including HIV/AIDS), or wound healing and  
CC disorders of epithelial cell proliferation. The nucleic acids are also  
CC useful for chromosome identification, radiation hybrid mapping or long-  
CC range restriction mapping, as molecular weight markers, or as  
CC hybridization or diagnostic probes. The polypeptides and antibodies are  
CC useful for providing immunological probes for differential identification  
CC of the tissues immunohistochemistry assays. Note: The sequence data for  
CC this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 427 BP; 115 A; 120 C; 102 G; 90 T; 0 U; 0 Other;

Query Match 98.4%; Score 101.4; DB 8; Length 427;  
Best Local Similarity 99.0%; Pred. No. 3.3e-27;  
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGACAGCGCAGCAGTCCAGCTCGGTTCTCTCGGTTTATCTGTAGAAATCAA 60

Db 281 CGGACAGCGCAGCAGTCCAGCTCGGTTCTCTCGGTTTATCTGTAGAAATCAA 340

Qy 61 ATGGTTCCTCCATAAATAAGGGCAGTGGCCCTTCTCTCAAAAAA 103

Db 341 ATGGTTCCTCCATAAATAAGGGCAGTGGCCCTTCTCTCAAAAAA 383

RESULT 7

ACH66744

ID ACH66744 standard; cDNA; 427 BP.

AC ACH66744;

XX 06-NOV-2003 (first entry)

XX Novel human secreted protein #97 cDNA.

XX Human; ss; vaccine; immune system disorder; haematopoietic cell disorder;  
KW cancer; autoimmune disorder; rheumatoid arthritis; glomerulonephritis;  
KW HIV infection; anaemia; thrombocytopenia blood coagulation disorder;  
KW blood platelet disorder; wound; heart attack; myocardial infarction;  
KW stroke; scarring; asthma; graft-versus host rejection; inflammation;  
KW hyperproliferative disorder; lymphoproliferative disorder; arrhythmia;  
KW aberrant cellular division; cell proliferative disorder; angiogenesis;  
KW cardiovascular disorder; pulmonary heart disease; neovascularisation;  
KW hypertrophic scar; keloid; ocular disorder; diabetic retinopathy; gene;  
KW uveitis; epithelial cell proliferation; neurological disease; apoptosis;  
KW Parkinson's disease; Alzheimer's disease; Huntington's chorea; AIDS;  
KW amyotrophic lateral sclerosis; toxin induced liver disease; septic shock;  
KW cachexia; anorexia; lung damage; infection.

XX Homo sapiens.

OS US2003065151-A1.

XX 03-APR-2003.

XX 04-APR-2002; 2002US-00115123.

XX 16-JUN-1998; 98US-0089507P.

XX 16-JUN-1998; 98US-0089508P.

XX 16-JUN-1998; 98US-0089509P.

XX 16-JUN-1998; 98US-0089510P.

XX 22-JUN-1998; 98US-0090112P.

XX 22-JUN-1998; 98US-0090113P.

XX 15-JUN-1999; 99WO-US013418.

XX 14-DEC-1999; 99US-00461325.

XX (HUMA-) HUMAN GENOME SCI INC.

PI Ruben SM, Ni J, Rosen CA, Wei Y, Young P, Florence K, Soppet DR;  
PI Brewer LA, Endress GA, Carter KC, Mucenski M, Ebner R, Lafleur DW;  
XX Olsen H, Shi Y, Moore PA, Komatsoulis G;  
DR WPI; 2003-531736/50.  
XX P-PSDB; ABO53454.

XX Novel antibody that binds specifically to a HCRJQ69 protein, useful for  
PT detecting the presence of a protein in a biological sample, and for  
PT treating cancers, autoimmune disorders and HIV infection.

XX Disclosure; SEQ ID NO 107; 176bp; English.

XX The invention relates to an isolated antibody or its fragment that  
CC specifically binds to a protein. The antibody is useful for detecting a  
CC protein in a biological sample, by contacting the biological sample with  
CC the antibody or its fragment and detecting the protein in the biological  
CC sample. The antibody is useful for purifying, detecting and targeting the  
CC human secreted proteins, including both in vitro and in vivo diagnostic  
CC and therapeutic methods. The antibody is useful for immunophenotyping of  
CC cell lines in biological samples and in antibody-based therapies for  
CC treating, inhibiting and preventing diseases, disorders or conditions  
CC associated with aberrant expression and/or activity of the above  
CC proteins. The antibody is useful for treating deficiencies or disorders  
CC of immune system and haematopoietic cells, for increasing differentiation  
CC and proliferation of haematopoietic cells, for treating immune  
CC rheumatoid arthritis and glomerulonephritis), HIV infection, anaemia and  
CC thrombocytopenia and as a marker for a particular immune system disease  
CC or disorder. The antibody is also useful for treating blood coagulation  
CC disorders, blood platelet disorders, wounds, heart attacks (infarction),  
CC strokes, scarring and asthma. The antibody is also useful for treating or  
CC preventing graft-versus host rejection, for modulating inflammation, for  
CC treating hyperproliferative disorders e.g. lymphoproliferative disorders  
CC and cancers, for inhibiting aberrant cellular division and for treating  
CC cell proliferative disorders. The antibody is also useful for treating  
CC cardiovascular disorders e.g. pulmonary heart disease and arrhythmia,  
CC disorders associated with neovascularisation and angiogenesis, for  
CC treating hypertrophic scars and keloids, ocular disorders e.g. diabetic  
CC retinopathy and uveitis, for wound healing and disorders of epithelial  
CC cell proliferation. The antibody is also useful for treating neurological  
CC diseases e.g. Parkinson's disease, Alzheimer's disease, Huntington's  
CC chorea and amyotrophic lateral sclerosis (ALS), diseases associated with  
CC increased apoptosis e.g. AIDS, toxin induced liver disease, septic shock,  
CC cachexia and anorexia, for preventing and healing damage to lungs and for  
CC treating infectious diseases. The present sequence represents cDNA  
CC encoding a novel human secreted protein. Note: the sequence data for this  
CC patent did not form part of the printed specification but was obtained in  
CC electronic format directly from USPTO at  
CC seqdata.uspto.gov/sequence.html?docID=20030065151

XX Sequence 427 BP; 115 A; 120 C; 102 G; 90 T; 0 U; 0 Other;

Query Match 98.4%; Score 101.4; DB 9; Length 427;  
Best Local Similarity 99.0%; Pred. No. 3.3e-27;  
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGACAGCGCAGCAGTCCAGCTCGGTTCTCTCGGTTTATCTGTAGAAATCAA 60

Db 281 CGGACAGCGCAGCAGTCCAGCTCGGTTCTCTCGGTTTATCTGTAGAAATCAA 340

Qy 61 ATGGTTCCTCCATAAATAAGGGCAGTGGCCCTTCTCTCAAAAAA 103

Db 341 ATGGTTCCTCCATAAATAAGGGCAGTGGCCCTTCTCTCAAAAAA 383

RESULT 8

ADB91346

ID ADB91346 standard; cDNA; 427 BP.

XX ADB91346;

XX 04-DEC-2003 (first entry)

XX Human secreted protein cDNA #SEQ ID 292.  
 XX Secreted protein; gene therapy; antidiabetic; diabetes; human; ss.  
 XX Homo sapiens.  
 XX W02003004622-A2.  
 XX 16-JAN-2003.  
 XX 19-MAR-2002; 2002WO-US008124.  
 XX 21-MAR-2001; 2001US-0277340P.  
 XX 19-JUL-2001; 2001US-0306171P.  
 XX 13-NOV-2001; 2001US-0331287P.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX Rosen CA, Ruben SM;  
 XX WPI; 2003-229407/22.  
 XX Nucleic acid encoding a human secreted protein is useful in diagnosing or  
 XX treating diabetes or conditions related to diabetes.  
 XX Claim 9; SEQ ID NO 292; 1537pp; English.  
 XX The invention relates to isolated nucleic acid molecules ADB91065-  
 XX ADB91448 and ADB91835-ADB91911 encoding human secreted proteins ADB91449-  
 XX ADB91834. Also disclosed is a recombinant vector comprising a  
 XX polynucleotide of the invention, and a recombinant host cell comprising  
 XX the recombinant vector. The polypeptide of the invention is useful in  
 XX identifying a binding partner by contacting the polypeptide with a  
 XX binding partner, and determining whether the binding partner increases or  
 XX decreases activity of the polypeptide. The polypeptide, polynucleotide,  
 XX antibody or its fragment, agonist or antagonist are useful for preparing  
 XX a pharmaceutical composition for diagnosing or treating diabetes or  
 XX conditions related to diabetes. The present sequence is that of the human  
 XX immunoglobulin Fc portion used to generate fusion proteins, increasing  
 XX the stability of the fused protein as compared to the secreted protein  
 XX only. Note: The sequence data for this patent did not form part of the  
 XX printed specification, but was obtained in electronic format directly  
 XX from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX Sequence 427 BP; 115 A; 120 C; 102 G; 90 T; 0 U; 0 Other;  
 XX Query Match 98.4%; Score 101.4; DB 9; Length 427;  
 XX Best Local Similarity 99.0%; Pred. No. 3.3e-27;  
 XX Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 CGGACAGCGCAGCAGCTCCAGCTCTGGTTCTCTCGGTTTATTCTGTAGATGAA 60  
 Db 281 CGGACAGCGCAGCAGCTCCAGCTCTGGTTCTCTCGGTTTATTCTGTAGATGAA 340  
 Qy 61 ATGGTTCCTCCATTAATAGGGGCATGACCCCTCTCTCAAAAA 103  
 Db 341 ATGGTTCCTCCATTAATAGGGGCATGACCCCTCTCTCAAAAA 383  
 RESULT 9  
 AAD34144  
 ID AAD34144 standard; cDNA; 439 BP.  
 XX  
 AC AAD34144;  
 XX  
 XX 16-JUL-2002 (first entry)  
 DT  
 XX Human secreted protein-encoding gene 17 cDNA clone HNH023, SEQ ID NO:49.  
 DE  
 XX Human; secreted protein; proliferative disorder; cancer; tumour; AIDS;  
 XX foetal abnormality; developmental; haematopoietic disorder; kidney; skin;  
 KW immune system; cardiovascular; angiogenic; gastrointestinal; endocrine;  
 KW

KW pregnancy-related disorder; cognitive; neurological; autoimmune disease;  
 KW inflammation; allergy; rheumatoid arthritis; Alzheimer's disease; asthma;  
 KW Parkinson's disease; schizophrenia; psoriasis; sepsis; atherosclerosis;  
 KW diabetes; infection; wound healing; vulnery; chemotaxis; food additive;  
 KW gene therapy; gene; ss.  
 XX Homo sapiens.  
 XX Key Location/Qualifiers  
 FH 12..245  
 CDS  
 FT /\*tag= a  
 FT /product= "Human secreted protein"  
 FT /note= "CDS does not include start codon"  
 FT /partial  
 FT sig\_peptide  
 FT 12..71  
 FT /\*tag= b  
 FT mat\_peptide  
 FT 72..242  
 FT /\*tag= c  
 FT /product= "Mature secreted protein"  
 XX  
 PN W0200222654-A1.  
 XX  
 XX 21-MAR-2002.  
 XX 17-JAN-2001; 2001WO-US001385.  
 XX 12-SEP-2000; 2000US-0231969P.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX Rosen CA, Komatsu GA, Baker KP, Birse CE, Soppet DR;  
 PI Olsen HS, Moore PA, Wel P, Ebner R, Duan DR, Shi Y, Choi GH;  
 PI Fiscella M, Ni J;  
 XX WPI; 2002-315684/35.  
 DR P-PSDB; AAE21616.  
 XX Seventeen nucleic acid molecules encoding human secreted proteins, useful  
 XX for treating and preventing cancer, immune disorders (e.g. Addison's  
 XX disease, and allergies), and cardiovascular disorders (e.g. myocardial  
 XX ischemias).  
 XX Claim 1; Page 436-437; 483pp; English.  
 XX AAD34106-AAD34144 represent cDNAs corresponding to 17 human secreted  
 XX protein genes, and AAE21578-AAE21616 represent the proteins they encode.  
 XX AAE21617-AAE21627 represent human secreted protein fragments. The genes  
 XX and their corresponding secreted proteins are useful for preventing,  
 XX treating or ameliorating medical conditions, e.g., by protein or gene  
 XX therapy. Pathological conditions can be diagnosed by determining the  
 XX amount of the new protein in a sample or by determining the presence of  
 XX mutations in the new genes. Specific uses are described for each of the  
 XX 17 genes, based on the tissues in which they are most highly expressed,  
 XX and include developing products for the diagnosis or treatment of cancer,  
 XX proliferative disorders, tumours, foetal and developmental abnormalities,  
 XX haematopoietic disorders, diseases of the immune system, AIDS, autoimmune  
 XX diseases (rheumatoid arthritis), inflammation, allergies, neurological  
 XX disorders (Alzheimer's disease, Parkinson's disease), cognitive disorder,  
 XX schizophrenia, asthma, skin disorders (psoriasis), sepsis, diabetes,  
 XX atherosclerosis, cardiovascular disorders, angiogenic disorders, kidney  
 XX disorders, gastrointestinal disorders, pregnancy-related disorders,  
 XX endocrine disorders, and infections. The proteins can also be used to aid  
 XX wound healing and epithelial cell proliferation, to prevent skin aging  
 XX due to sunburn, to maintain organs before transplantation, for supporting  
 XX cell culture of primary tissues, to regenerate tissues, to identify their  
 XX cognate ligands or binding partners, and in chemotaxis, and can be used  
 XX as a food additive or preservative to modify storage properties. The  
 XX present sequence represents a human secreted protein-encoding cDNA of the  
 XX invention  
 XX  
 SQ Sequence 439 BP; 124 A; 121 C; 102 G; 90 T; 0 U; 2 Other;  
 XX Query Match 98.4%; Score 101.4; DB 6; Length 439;

Best Local Similarity 99.0%; Pred. No. 3.3e-27;		Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
Qy	1	CGGACAGACGGCAGCAGTCCAGCTCTGGTTTCCTTCCTCGGTTTATTCGTGTAGAA	60
	281	CGGACAGACGGCAGCAGTCCAGCTCTGGTTTCCTTCCTCGGTTTATTCGTGTAGAA	340
Db	61	ATGGTTCCTCCATAAATAAGGGGCATGAGCCCTTCCTCACA	103
Qy	341	ATGGTTCCTCCATAAATAAGGGGCATGAGCCCTTCCTCACA	383
Db			
RESULT 10			
AAK56586/c			
ID	AAK56586 standard; cDNA; 1148 BP.		
XX			
AC	AAK56586;		
XX			
DT	06-NOV-2001 (first entry)		
XX			
DE	Human immune/haematopoietic antigen encoding cDNA SEQ ID NO:1646.		
XX			
KW	Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;		
KW	cytostatic; gene therapy; vaccine; metastasis; ss.		
OS	Homo sapiens.		
XX			
PN	WO200157182-A2.		
XX			
PD	09-AUG-2001.		
XX			
PF	17-JAN-2001; 2001WO-US001354.		
XX			
PR	31-JAN-2000; 2000US-0179065P.		
PR	04-FEB-2000; 2000US-0180628P.		
PR	24-FEB-2000; 2000US-0184664P.		
PR	02-MAR-2000; 2000US-0186350P.		
PR	16-MAR-2000; 2000US-0189874P.		
PR	17-MAR-2000; 2000US-0190076P.		
PR	18-APR-2000; 2000US-0198123P.		
PR	19-MAY-2000; 2000US-0205515P.		
PR	07-JUN-2000; 2000US-0209467P.		
PR	28-JUN-2000; 2000US-0214886P.		
PR	30-JUN-2000; 2000US-0215135P.		
PR	07-JUL-2000; 2000US-0216647P.		
PR	07-JUL-2000; 2000US-0216880P.		
PR	11-JUL-2000; 2000US-0217487P.		
PR	11-JUL-2000; 2000US-0217496P.		
PR	14-JUL-2000; 2000US-0218290P.		
PR	26-JUL-2000; 2000US-0220963P.		
PR	26-JUL-2000; 2000US-0220964P.		
PR	14-AUG-2000; 2000US-0224518P.		
PR	14-AUG-2000; 2000US-0224519P.		
PR	14-AUG-2000; 2000US-0225213P.		
PR	14-AUG-2000; 2000US-0225214P.		
PR	14-AUG-2000; 2000US-0225266P.		
PR	14-AUG-2000; 2000US-0225267P.		
PR	14-AUG-2000; 2000US-0225268P.		
PR	14-AUG-2000; 2000US-0225270P.		
PR	14-AUG-2000; 2000US-0225447P.		
PR	14-AUG-2000; 2000US-0225757P.		
PR	14-AUG-2000; 2000US-0225758P.		
PR	14-AUG-2000; 2000US-0225759P.		
PR	18-AUG-2000; 2000US-0226279P.		
PR	22-AUG-2000; 2000US-0226681P.		
PR	22-AUG-2000; 2000US-0226686P.		
PR	22-AUG-2000; 2000US-0227182P.		
PR	23-AUG-2000; 2000US-0227009P.		
PR	30-AUG-2000; 2000US-0228924P.		
PR	01-SEP-2000; 2000US-0229287P.		
PR	01-SEP-2000; 2000US-0229343P.		
PR	01-SEP-2000; 2000US-0229344P.		
PR	01-SEP-2000; 2000US-0229345P.		

PR	05-SEP-2000; 2000US-0229509P.
PR	05-SEP-2000; 2000US-0229513P.
PR	06-SEP-2000; 2000US-0230437P.
PR	06-SEP-2000; 2000US-0230438P.
PR	08-SEP-2000; 2000US-0231242P.
PR	08-SEP-2000; 2000US-0231243P.
PR	08-SEP-2000; 2000US-0231244P.
PR	08-SEP-2000; 2000US-0231413P.
PR	08-SEP-2000; 2000US-0231414P.
PR	08-SEP-2000; 2000US-0232080P.
PR	08-SEP-2000; 2000US-0232081P.
PR	12-SEP-2000; 2000US-0231988P.
PR	14-SEP-2000; 2000US-0232397P.
PR	14-SEP-2000; 2000US-0232398P.
PR	14-SEP-2000; 2000US-0232399P.
PR	14-SEP-2000; 2000US-0232400P.
PR	14-SEP-2000; 2000US-0232401P.
PR	14-SEP-2000; 2000US-0233063P.
PR	14-SEP-2000; 2000US-0233064P.
PR	21-SEP-2000; 2000US-0234223P.
PR	21-SEP-2000; 2000US-0234274P.
PR	25-SEP-2000; 2000US-0234997P.
PR	25-SEP-2000; 2000US-0234998P.
PR	26-SEP-2000; 2000US-0235484P.
PR	27-SEP-2000; 2000US-0235834P.
PR	27-SEP-2000; 2000US-0235836P.
PR	29-SEP-2000; 2000US-0236327P.
PR	29-SEP-2000; 2000US-0236368P.
PR	29-SEP-2000; 2000US-0236369P.
PR	29-SEP-2000; 2000US-0236370P.
PR	02-OCT-2000; 2000US-0236802P.
PR	02-OCT-2000; 2000US-0237037P.
PR	02-OCT-2000; 2000US-0237038P.
PR	02-OCT-2000; 2000US-0237039P.
PR	13-OCT-2000; 2000US-0239935P.
PR	13-OCT-2000; 2000US-0239937P.
PR	20-OCT-2000; 2000US-0240960P.
PR	20-OCT-2000; 2000US-0241221P.
PR	20-OCT-2000; 2000US-0241785P.
PR	20-OCT-2000; 2000US-0241786P.
PR	20-OCT-2000; 2000US-0241787P.
PR	20-OCT-2000; 2000US-0241808P.
PR	20-OCT-2000; 2000US-0241809P.
PR	20-OCT-2000; 2000US-0241826P.
PR	01-NOV-2000; 2000US-0244617P.
PR	08-NOV-2000; 2000US-0246474P.
PR	08-NOV-2000; 2000US-0246475P.
PR	08-NOV-2000; 2000US-0246476P.
PR	08-NOV-2000; 2000US-0246477P.
PR	08-NOV-2000; 2000US-0246478P.
PR	08-NOV-2000; 2000US-0246523P.
PR	08-NOV-2000; 2000US-0246524P.
PR	08-NOV-2000; 2000US-0246525P.
PR	08-NOV-2000; 2000US-0246526P.
PR	08-NOV-2000; 2000US-0246527P.
PR	08-NOV-2000; 2000US-0246528P.
PR	08-NOV-2000; 2000US-0246532P.
PR	08-NOV-2000; 2000US-0246609P.
PR	08-NOV-2000; 2000US-0246610P.
PR	08-NOV-2000; 2000US-0246611P.
PR	17-NOV-2000; 2000US-0249207P.
PR	17-NOV-2000; 2000US-0249208P.
PR	17-NOV-2000; 2000US-0249209P.
PR	17-NOV-2000; 2000US-0249210P.
PR	17-NOV-2000; 2000US-0249211P.
PR	17-NOV-2000; 2000US-0249212P.
PR	17-NOV-2000; 2000US-0249213P.
PR	17-NOV-2000; 2000US-0249214P.
PR	17-NOV-2000; 2000US-0249215P.







Db 1974 ATGGTCCCATATAAAGGGGCGATGAGCCCTTCTCAGCAAAA 2016

RESULT 14

AAI70679

ID AAI70679 standard; cDNA; 2060 BP.

AC AAI70679;

XX

XX 04-FEB-2002 (first entry)

DT Human membrane transporter (MTP) 33556 cDNA clone Fbh33556.

DE

XX

XX Membrane transporter; MTP; human; cell proliferation; cell growth;

XX cell differentiation; cell migration; immune response; hormone response;

XX cell communication; nootropic; neuroprotective; anticonvulsant;

XX hypotensive; antidepressant; neuroleptic; antiarteriosclerotic;

XX vasotropic; antianginal; cardiant; thrombolytic; cytostatic;

XX antiinfertility; immunosuppressive; immunostimulant;

XX proton-dependent oligopeptide transport; POT; diagnosis; gene therapy;

XX ss.

XX

XX Homo sapiens.

XX

XX Key Location/Qualifiers

PH 136..1746

FT /\*tag= a

FT /note= "the coding region is also claimed in Claim 1(a)"

XX

XX W0200179497-A2.

XX

XX 25-OCT-2001.

XX

XX 13-APR-2001; 2001WO-US012187.

XX

XX 14-APR-2000; 2000US-0197376P.

XX

XX (MTLL-) MILLENNIUM PHARM INC.

XX

XX Curtis RAJ;

XX

XX WPI; 2002-034357/04.

XX

XX P-PSDB; AAM50318.

XX

XX Novel 33556 (a membrane transporter) protein and nucleic acids encoding

XX the protein, useful for treating disorders associated with transporter-

XX mediated activity such as Alzheimer's disease, arteriosclerosis, cancer.

XX

XX Claim 1(a); Fig 1A-D; 115pp; English.

XX

XX The present sequence, the coding region of which is also claimed, is that

XX of cDNA clone Fbh33556, which codes for novel 581-amino acid (64 kDa)

XX human membrane transporter (MTP) protein 33556 (see AAM50318). MTP is a

XX novel member of the proton-dependent oligopeptide transport (POT) family

XX of proteins possessing the ability to shuttle molecules across a lipid

XX bilayer. MTP proteins are capable of transporting ions, proteins, and

XX small molecules across biological membranes both within a cell and

XX between the cell and the environment and, thus, play a role in or

XX function in a variety of cellular processes, e.g. proliferation, growth,

XX differentiation, migration, immune responses, hormonal responses, and

XX inter- or intracellular communication. The invention provides MTP nucleic

XX acids, antisense nucleic acids, recombinant expression vectors, host

XX cells and nonhuman transgenic animals in which an MTP gene has been

XX introduced or disrupted, as well as MTP proteins, antigenic peptides and

XX anti-MTP antibodies. These are used in claimed methods for identifying a

XX subject at risk of developing an MTP disorder, and for identifying a

XX modulator compounds capable of treating such as disorder. The 33556

XX modulator may be an antisense 33556 nucleic acid, a ribozyme, the present

XX nucleic acid sequence or its fragment or variant. 33556 Nucleic acids are

XX also used in the recombinant production of 33556 polypeptides, and as

XX probes and primers. Examples of MTP-associated or related disorders

XX include: transporter-associated disorders such as CNS disorders including

XX Alzheimer's disease, senile dementia, Huntington's disease, multiple

CC sclerosis, amyotrophic lateral sclerosis, Jakob-Creutzfeldt disease,

CC autonomic function disorders such as hypertension and sleep disorders,

CC and neuropsychiatric disorders such as depression, schizophrenia,

CC learning or memory disorders such as amnesia or age-related memory loss,

CC attention deficit disorder, dysthymic disorder, depression, bipolar

CC obsessive compulsive disorder, anxiety, phobia, panic, and manic

CC affective disorder; cardiac-related disorders such as arteriosclerosis,

CC ischaemic reperfusion injury, restenosis, arterial inflammation, vascular

CC wall remodeling, coronary microembolism, tachycardia, bradycardia,

CC anginal, hypertension, idiopathic cardiomyopathy, myocardial infarction;

CC cancer such as carcinoma, sarcoma or leukaemia, metastasis, skeletal

CC dysplasia, hepatic disorders and haematopoietic and/or myeloproliferative

CC disorders; hormonal disorders such as type I and type II diabetes,

CC pituitary disorders, thyroid disorders, reproductive or fertility

CC disorders; autoimmune or immune deficiency disorders, such as congenital

CC X-linked infantile hypogammaglobulinaemia, transient

CC hypogammaglobulinaemia, selective IgA deficiency; and muscular disorders

CC such as paralysis, muscle weakness, muscular dystrophy, congenital

CC myopathies. The 33556 molecules are also useful as markers of disorders

CC or disease states, as markers for precursors of disease states, as

CC markers for predisposition of disease states, as markers of drug

CC activity, or as markers of the pharmacogenomic profile of a subject

XX

XX

SQ Sequence 2060 BP; 330 A; 694 C; 613 G; 415 T; 0 U; 8 Other;

Query Match 98.4%; Score 101.4; DB 6; Length 2060;

Best Local Similarity 99.0%; Pred. No. 6.2e-27;

Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGACAGACGGCAGCAGTCCAGCTCTGGTTTCTCTCTCGGTTTATTCTGTAGATGAA 60

DB 1917 CGGACAGACGGCAGCAGTCCAGCTCTGGTTTCTCTCTCGGTTTATTCTGTAGATGAA 1976

QY 61 ATGGTTCCCATATAAATAAGGGGCGATGAGCCCTTCTCAGCAAAA 103

DB 1977 ATGGTTCCCATATAAATAAGGGGCGATGAGCCCTTCTCAGCAAAA 2019

RESULT 15

AAK52347

ID AAK52347 standard; cDNA; 2124 BP.

XX

XX AAK52347;

XX

XX 06-NOV-2001 (first entry)

DT Human polynucleotide SEQ ID NO 892.

DE

XX

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;

XX vaccine; peptide therapy; stem cell growth factor; haematopoiesis;

XX tissue growth factor; immunomodulatory; cancer; leukaemia;

XX nervous system disorder; arthritis; inflammation; ss.

XX

XX Homo sapiens.

XX

XX W0200157190-A2.

XX

XX 09-AUG-2001.

XX

XX 05-FEB-2001; 2001WO-US0004098.

XX

XX 03-FEB-2000; 2000US-00496914.

XX 27-APR-2000; 2000US-00560875.

XX 20-JUN-2000; 2000US-00598075.

XX 19-JUL-2000; 2000US-00620325.

XX 01-SEP-2000; 2000US-00654936.

XX 15-SEP-2000; 2000US-00663561.

XX 20-OCT-2000; 2000US-00693325.

XX 30-NOV-2000; 2000US-00728422.

XX

XX (HYSE-) HYSEQ INC.

XX

XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;

PI



PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;  
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;  
 XX  
 DR WPI: 2001-476283/51.  
 DR P-PSDB; AAM79214.  
 XX  
 PT Nucleic acids encoding polypeptides with cytokine-like activities, useful  
 PT in diagnosis and gene therapy.  
 XX  
 PS Claim 1; Page 2937-2939; 6221pp; English.  
 XX  
 CC The invention relates to polynucleotides (AAK51456-AAK53435) and the  
 CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to  
 CC cytokine, cell proliferation or cell differentiation or which may induce  
 CC production of other cytokines in other cell populations. The  
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
 CC peptide therapy. The polypeptides have various cytokine-like activities,  
 CC e.g. stem cell growth factor activity, haematopoiesis regulating  
 CC activity, tissue growth factor activity, immunomodulatory activity and  
 CC activin/inhibin activity and may be useful in the diagnosis and/or  
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
 CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111  
 CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the  
 CC sequence listing were missing at the time of publication  
 XX  
 SQ Sequence 2124 BP; 347 A; 691 C; 666 G; 420 T; 0 U; 0 Other;

Query Match 98.4%; Score 101.4; DB 4; Length 2124;  
 Best Local Similarity 99.0%; Pred. No. 6.3e-27;  
 Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGACAGAGCGCAGCAGTCCAGCTCTGGTTTCCTTCCTTCGCTTTATCTCTGTAAGTAA 60  
 Db 2016 CGGACAGAGCGCAGCAGTCCAGCTCTGGTTTCCTTCCTTCGCTTTATCTCTGTAAGTAA 2075  
 |||||||  
 Qy 61 ATGGTTCCCAATAAAGGGGCATGAGCCCTTCCTCAAAAAA 103  
 Db 2076 ATGGTTCCCAATAAAGGGGCATGAGCCCTTCCTCAAAAAA 2118  
 |||||||

Search completed: December 7, 2005, 00:03:12  
 Job time : 163.924 secs

*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

# OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 11:31:40 ; Search time 256.473 Seconds  
(without alignments)  
3320.997 Million cell updates/sec

Title: US-09-980-046B-8

Perfect score: 103

Sequence: 1 cggacagcggcagcagtc.....tgagcccttcctcacaaaa 103

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Database :

Published Applications NA\_Main:  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US11\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	102.2	99.2	1149	6	US-10-264-237-500
2	101.4	98.4	427	5	US-10-012-542-107
3	101.4	98.4	427	5	US-10-115-123-107
4	101.4	98.4	427	7	US-10-800-834-107
5	101.4	98.4	1874	8	US-10-775-920-276
6	101.4	98.4	2034	8	US-10-775-920-277
7	101.4	98.4	2057	3	US-09-835-270-1
8	101.4	98.4	2057	3	US-10-753-267-67
9	100.4	97.5	1277	5	US-10-012-542-20
10	100.4	97.5	1277	5	US-10-115-123-20
11	100.4	97.5	1277	7	US-10-800-834-20
12	98.4	95.5	26555	3	US-09-860-670-161
13	98.4	95.5	26555	6	US-10-227-646-161
14	98.4	95.5	136328	6	US-10-101-510-127
15	96.4	93.6	2113	6	US-10-172-118-1657
16	96.4	93.6	2113	7	US-10-342-887-1657
17	96.4	93.6	2113	8	US-10-755-889-203
18	96.4	93.6	2113	8	US-10-775-920-274
19	96.4	93.6	2113	8	US-10-775-920-275
20	90.4	87.8	1104	7	US-10-276-774-720
21	74.4	72.2	280	3	US-09-745-288-58
22	74.4	72.2	280	7	US-10-453-919-58
23	74.4	72.2	2786	3	US-09-814-353-20133

c 24 72.8 70.7 2755 6 US-10-269-909-32 Sequence 32, Appl  
c 25 72.8 70.7 2755 8 US-10-723-860-1705 Sequence 1705, Ap  
c 26 72.8 70.7 2755 9 US-10-936-626-33 Sequence 33, Appl  
c 27 72.8 70.7 2755 9 US-10-938-061-33 Sequence 33, Appl  
c 28 72.8 70.7 2769 8 US-10-723-860-6089 Sequence 6089, Ap  
c 29 71.8 69.7 345 6 US-10-062-674-1501 Sequence 1501, Ap  
c 30 70 68.0 70 9 US-10-957-432-406 Sequence 406, App  
31 68 66.0 1902 6 US-10-257-022-26 Sequence 26, Appl  
32 63.8 61.9 317 3 US-09-968-007A-622 Sequence 622, App  
33 63.8 61.9 317 3 US-09-968-007A-914 Sequence 914, App  
34 63.8 61.9 317 9 US-10-843-641A-7092 Sequence 7092, Ap  
35 63.8 61.9 317 9 US-10-843-641A-7384 Sequence 7384, Ap  
36 43 41.7 281 3 US-09-960-352-1755 Sequence 1755, Ap  
c 37 40.2 39.0 346 6 US-10-062-674-362 Sequence 362, App  
c 38 33.4 32.4 1132 4 US-09-925-065A-92388 Sequence 92388, A  
39 32.2 31.3 594 3 US-09-764-891-536 Sequence 536, App  
40 31 30.1 1401 7 US-10-282-122A-17398 Sequence 17398, A  
c 41 28.8 28.0 222 3 US-09-783-590-636 Sequence 636, App  
c 42 28.8 28.0 553 4 US-09-925-065A-517586 Sequence 517586, A  
43 28.8 28.0 1284 7 US-10-425-114-10355 Sequence 10355, A  
44 28.8 28.0 2366 7 US-10-424-599-39215 Sequence 39215, A  
45 28.8 28.0 2672 7 US-10-424-599-39212 Sequence 39212, A

## ALIGNMENTS

RESULT 1  
US-10-264-237-500  
; Sequence 500, Application US/10264237  
; Publication No. US20040009491A1  
; GENERAL INFORMATION:  
; APPLICANT: Birese et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: P4131P1  
; CURRENT APPLICATION NUMBER: US/10/264,237  
; CURRENT FILING DATE: 2002-10-04  
; PRIOR APPLICATION NUMBER: PCT/US01/16450  
; PRIOR FILING DATE: 2001-05-18  
; PRIOR APPLICATION NUMBER: US 60/205,515  
; PRIOR FILING DATE: 2000-05-19  
; NUMBER OF SEQ ID NOS: 2876  
; SOFTWARE: PatentIn Ver. 3.1  
; SEQ ID NO 500  
; LENGTH: 1149  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (99)..(99)  
; OTHER INFORMATION: n equals a,t,g, or c  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1133)..(1134)  
; OTHER INFORMATION: n equals a,t,g, or c  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1144)..(1144)  
; OTHER INFORMATION: n equals a,t,g, or c  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1149)..(1149)  
; OTHER INFORMATION: n equals a,t,g, or c  
US-10-264-237-500

Query Match 99.2%; Score 102.2; DB 6; Length 1149;

Best Local Similarity 98.1%; Pred. No. 1.9e-26;

Matches 101; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGACAGCGCAGCAGCTCCAGCTCTCGGTTTCCTTCGTTTATTCGTAGATGAA 60

DB 991 CGGACAGCGCAGCAGCTCTCGGTTTCCTTCGTTTATTCGTAGATGAA 1050

QY 61 ATGTTCCCATTAATAAGGGCGCATGAGCCCTTCTCACAATAA 103  
|||||  
Db 1051 ATGTTCCCATTAATAAGGGCGCATGAGCCCTTCTCACAATAA 1093  
|||||

## RESULT 2

US-10-012-542-107  
; Sequence 107, Application US/10012542  
; Publication No. US20030044851A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: 94 Human Secreted Proteins  
; FILE REFERENCE: P2029P1  
; CURRENT APPLICATION NUMBER: US/10/012,542  
; CURRENT FILING DATE: 2001-12-12  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/461,325  
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-12-14  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,507  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,508  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,509  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,510  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,112  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,113  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22  
; NUMBER OF SEQ ID NOS: 532  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 107  
; LENGTH: 427  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-012-542-107

Query Match 98.4%; Score 101.4; DB 5; Length 427;  
Best Local Similarity 99.0%; Pred. No. 2.5e-26;  
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CGGACAGACGGCAGCAGTCCAGCTCTGGTTTCTCTCGGTTTATCTCTTGAATGAA 60  
|||||  
Db 281 CGGACAGACGGCAGCAGTCCAGCTCTGGTTTCTCTCGGTTTATCTCTTGAATGAA 340  
|||||  
QY 61 ATGTTCCCATTAATAAGGGCGCATGAGCCCTTCTCACAATAA 103  
|||||  
Db 341 ATGTTCCCATTAATAAGGGCGCATGAGCCCTTCTCACAATAA 383  
|||||

## RESULT 3

US-10-115-123-107  
; Sequence 107, Application US/10115123  
; Publication No. US20030065151A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: 94 Human Secreted Proteins  
; FILE REFERENCE: P2029G3OAP1D2  
; CURRENT APPLICATION NUMBER: US/10/115,123  
; CURRENT FILING DATE: 2002-04-04  
; PRIOR APPLICATION NUMBER: PCT/US99/13418  
; PRIOR FILING DATE: 1999-06-15  
; PRIOR APPLICATION NUMBER: 60/089,507  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089,508  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089,509  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089,510  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/090,112  
; PRIOR FILING DATE: 1998-06-22  
; PRIOR APPLICATION NUMBER: 60/090,113

; PRIOR FILING DATE: 1998-06-22  
; NUMBER OF SEQ ID NOS: 532  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 107  
; LENGTH: 427  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-115-123-107

Query Match 98.4%; Score 101.4; DB 5; Length 427;  
Best Local Similarity 99.0%; Pred. No. 2.5e-26;  
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CGGACAGACGGCAGCAGTCCAGCTCTGGTTTCTCTCGGTTTATCTCTTGAATGAA 60  
|||||  
Db 281 CGGACAGACGGCAGCAGTCCAGCTCTGGTTTCTCTCGGTTTATCTCTTGAATGAA 340  
|||||  
QY 61 ATGTTCCCATTAATAAGGGCGCATGAGCCCTTCTCACAATAA 103  
|||||  
Db 341 ATGTTCCCATTAATAAGGGCGCATGAGCCCTTCTCACAATAA 383  
|||||

## RESULT 4

US-10-800-834-107  
; Sequence 107, Application US/10800834  
; Publication No. US20040146930A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: 94 Human Secreted Proteins  
; FILE REFERENCE: P2029P1D3  
; CURRENT APPLICATION NUMBER: US/10/800,834  
; CURRENT FILING DATE: 2004-03-16  
; PRIOR APPLICATION NUMBER: 10/115,123  
; PRIOR FILING DATE: 2002-04-04  
; PRIOR APPLICATION NUMBER: 09/461,325  
; PRIOR FILING DATE: 1999-12-14  
; PRIOR APPLICATION NUMBER: PCT/US99/13418  
; PRIOR FILING DATE: 1999-06-15  
; PRIOR APPLICATION NUMBER: 60/089,507  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089,508  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089,509  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089,510  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/090,112  
; PRIOR FILING DATE: 1998-06-22  
; PRIOR APPLICATION NUMBER: 60/090,113  
; PRIOR FILING DATE: 1998-06-22  
; NUMBER OF SEQ ID NOS: 532  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 107  
; LENGTH: 427  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-800-834-107

Query Match 98.4%; Score 101.4; DB 7; Length 427;  
Best Local Similarity 99.0%; Pred. No. 2.5e-26;  
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CGGACAGACGGCAGCAGTCCAGCTCTGGTTTCTCTCGGTTTATCTCTTGAATGAA 60  
|||||  
Db 281 CGGACAGACGGCAGCAGTCCAGCTCTGGTTTCTCTCGGTTTATCTCTTGAATGAA 340  
|||||  
QY 61 ATGTTCCCATTAATAAGGGCGCATGAGCCCTTCTCACAATAA 103  
|||||  
Db 341 ATGTTCCCATTAATAAGGGCGCATGAGCCCTTCTCACAATAA 383  
|||||

## RESULT 5

US-10-775-920-276

```
; Sequence 276, Application US/10775920
; Publication No. US2004017544A1
; GENERAL INFORMATION:
; APPLICANT: Mergen Ltd
; TITLE OF INVENTION: POLYNUCLEOTIDE SEQUENCES AND CORRESPONDING ENCODED POLYPEPTIDES
; TITLE OF INVENTION: OF PARTICULAR SECRETED AND MEMBRANE-BOUND PROTEINS OVEREXPRESSED
; FILE REFERENCE: Mergen - 0010B
; CURRENT APPLICATION NUMBER: US/10/775,920
; CURRENT FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: US 60/447,900
; PRIOR FILING DATE: 2003-02-13
; NUMBER OF SEQ ID NOS: 385
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 276
; LENGTH: 1874
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-775-920-276

Query Match 98.4%; Score 101.4; DB 8; Length 1874;
Best Local Similarity 99.0%; Pred. No. 4.5e-26;
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGACAGACGGCAGCAGTCCAGCTCTGGTTTCCTCTCGGTTTATTCTGTTAGATGAA 60
Db 1707 CGGACAGACGGCAGCAGTCCAGCTCTGGTTTCCTCTCGGTTTATTCTGTTAGATGAA 1766

Qy 61 ATGGTTCCCATAAATAAGGGGCATGAGCCCTTCTCACAAAA 103
Db 1767 ATGGTTCCCATAAATAAGGGGCATGAGCCCTTCTCACAAAA 1809

RESULT 6
US-10-775-920-277
; Sequence 277, Application US/10775920
; Publication No. US2004017544A1
; GENERAL INFORMATION:
; APPLICANT: Mergen Ltd
; TITLE OF INVENTION: POLYNUCLEOTIDE SEQUENCES AND CORRESPONDING ENCODED POLYPEPTIDES
; TITLE OF INVENTION: OF PARTICULAR SECRETED AND MEMBRANE-BOUND PROTEINS OVEREXPRESSED
; FILE REFERENCE: Mergen - 0010B
; CURRENT APPLICATION NUMBER: US/10/775,920
; CURRENT FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: US 60/447,900
; PRIOR FILING DATE: 2003-02-13
; NUMBER OF SEQ ID NOS: 385
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 277
; LENGTH: 2034
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-775-920-277

Query Match 98.4%; Score 101.4; DB 8; Length 2034;
Best Local Similarity 99.0%; Pred. No. 4.7e-26;
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGACAGACGGCAGCAGTCCAGCTCTGGTTTCCTCTCGGTTTATTCTGTTAGATGAA 60
Db 1876 CGGACAGACGGCAGCAGTCCAGCTCTGGTTTCCTCTCGGTTTATTCTGTTAGATGAA 1935

Qy 61 ATGGTTCCCATAAATAAGGGGCATGAGCCCTTCTCACAAAA 103
Db 1936 ATGGTTCCCATAAATAAGGGGCATGAGCCCTTCTCACAAAA 1978

RESULT 7
US-09-835-270-1
; Sequence 1, Application US/09835270
; Patent No. US20020077462A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Curtis, Rory A.
; TITLE OF INVENTION: 33556, A NOVEL HUMAN TRANSPORTER AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: 381553000800
; CURRENT APPLICATION NUMBER: US/09/835,270
; CURRENT FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/197,376
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2057
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (133)...(1878)
; OTHER INFORMATION: CDNA of human MTP
US-09-835-270-1

Query Match 98.4%; Score 101.4; DB 3; Length 2057;
Best Local Similarity 99.0%; Pred. No. 4.7e-26;
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGACAGACGGCAGCAGTCCAGCTCTGGTTTCCTCTCGGTTTATTCTGTTAGATGAA 60
Db 1914 CGGACAGACGGCAGCAGTCCAGCTCTGGTTTCCTCTCGGTTTATTCTGTTAGATGAA 1973

Qy 61 ATGGTTCCCATAAATAAGGGGCATGAGCCCTTCTCACAAAA 103
Db 1974 ATGGTTCCCATAAATAAGGGGCATGAGCCCTTCTCACAAAA 2016

RESULT 8
US-10-753-267-67
; Sequence 67, Application US/10753267
; Publication No. US20050037946A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Stagliano, Nancy E.
; APPLICANT: Healy, Aileen L.
; APPLICANT: Acton, Susan L.
; APPLICANT: Galvin, Katherine M.
; APPLICANT: Donoghue, Mary A.
; APPLICANT: Rodrigue-Way, Amelie
; APPLICANT: Tomlinson, James E.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
; TITLE OF INVENTION: CARDIOVASCULAR DISEASE USING 1722, 10280, 59917, 85553,
; TITLE OF INVENTION: 10653, 9235, 21668, 17794, 2210, 6169, 10102, 21061, 17662,
; TITLE OF INVENTION: 1468, 12282, 6350, 9035, 1820, 23652, 7301, 8925, 8701,
; TITLE OF INVENTION: 3533, 9462, 9123, 12788, 17729, 65552, 1261, 21476, 33770,
; TITLE OF INVENTION: 9380, 2565654, 33556, 53656, 44143, 32612, 10671, 261,
; TITLE OF INVENTION: 44570, 41922, 2552, 2417, 19319, 43969, 8921, 8993, 955,
; TITLE OF INVENTION: 32345, 966, 1920, 17318, 1510, 14180, 26005, 554, 16408,
; TITLE OF INVENTION: 42028, 112091, 13886, 13942, 1673, 54946 OR 2419
; FILE REFERENCE: MPI03-003PIRNOVMIM
; CURRENT APPLICATION NUMBER: US/10/753,267
; CURRENT FILING DATE: 2004-01-08
; PRIOR APPLICATION NUMBER: US 60/439,683
; PRIOR FILING DATE: 2003-01-13
; PRIOR APPLICATION NUMBER: US 60/445,216
; PRIOR FILING DATE: 2003-02-05
; PRIOR APPLICATION NUMBER: US 60/448,036
; PRIOR FILING DATE: 2003-02-18
; PRIOR APPLICATION NUMBER: US 60/454,189
; PRIOR FILING DATE: 2003-03-12
; PRIOR APPLICATION NUMBER: US 60/457,541
; PRIOR FILING DATE: 2003-03-25
; PRIOR APPLICATION NUMBER: US 60/466,411
; PRIOR FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: US 60/469,041
; PRIOR FILING DATE: 2003-05-08
; PRIOR APPLICATION NUMBER: US 60/477,414
```

```
; PRIOR FILING DATE: 2003-06-10
; PRIOR APPLICATION NUMBER: US 60/478,560
; PRIOR FILING DATE: 2003-06-13
; PRIOR APPLICATION NUMBER: US 60/489,772
; PRIOR FILING DATE: 2003-07-24
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 67
; LENGTH: 2057
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (133)....(1878)
US-10-753-267-67

Query Match          98.4%; Score 101.4; DB 8; Length 2057;
Best Local Similarity 99.0%; Pred. No. 4.7e-26;
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1  CGACAGACGGCAGCAGTCCAGCTCTGGTTCTCTCGGTTTATTCTGTTAGAAATGAA 60
Db      1914 CGACAGACGGCAGCAGTCCAGCTCTGGTTCTCTCGGTTTATTCTGTTAGAAATGAA 1973
;
QY      61  ATGTTCCCATTAATAAGGGGCATGAGCCCTTCTCACAATAA 103
Db      1974 ATGTTCCCATTAATAAGGGGCATGAGCCCTTCTCACAATAA 2016
;
RESULT 9
US-10-012-542-20
; Sequence 20, Application US/10012542
; Publication No. US2003004851A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 94 Human Secreted Proteins
; FILE REFERENCE: P2029P1
; CURRENT APPLICATION NUMBER: US/10/012,542
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/461,325
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-12-14
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,507
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,508
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,509
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,510
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,112
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,113
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 1277
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1207)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1272)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-012-542-20

Query Match          97.5%; Score 100.4; DB 5; Length 1277;
Best Local Similarity 98.1%; Pred. No. 9e-26;
Matches 101; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1  CGACAGACGGCAGCAGTCCAGCTCTGGTTCTCTCGGTTTATTCTGTTAGAAATGAA 60
Db      1914 CGACAGACGGCAGCAGTCCAGCTCTGGTTCTCTCGGTTTATTCTGTTAGAAATGAA 1189
;
QY      61  ATGTTCCCATTAATAAGGGGCATGAGCCCTTCTCACAATAA 103
Db      1974 ATGTTCCCATTAATAAGGGGCATGAGCCCTTCTCACAATAA 1232
;
RESULT 11
US-10-800-834-20
; Sequence 20, Application US/10800834
; Publication No. US20040146930A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 94 Human Secreted Proteins
; FILE REFERENCE: P2029P1D3
; CURRENT APPLICATION NUMBER: US/10/800,834
; CURRENT FILING DATE: 2004-03-16
; PRIOR APPLICATION NUMBER: 10/115,123
; PRIOR FILING DATE: 2002-04-04
```

```
QY      1  CGACAGACGGCAGCAGTCCAGCTCTGGTTCTCTCGGTTTATTCTGTTAGAAATGAA 60
Db      1130 CGACAGACGGCAGCAGTCCAGCTCTGGTTCTCTCGGTTTATTCTGTTAGAAATGAA 1189
;
QY      61  ATGTTCCCATTAATAAGGGGCATGAGCCCTTCTCACAATAA 103
Db      1190 ATGTTCCCATTAATAAGGGGCATGAGCCCTTCTCACAATAA 1232
;
RESULT 10
US-10-115-123-20
; Sequence 20, Application US/10115123
; Publication No. US20030065151A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 94 Human Secreted Proteins
; FILE REFERENCE: P2029G3OAP1D2
; CURRENT APPLICATION NUMBER: US/10/115,123
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: PCT/US99/13418
; PRIOR FILING DATE: 1999-06-15
; PRIOR APPLICATION NUMBER: 60/089,507
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089,508
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089,509
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089,510
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/090,112
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090,113
; PRIOR FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 1277
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1207)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1272)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-115-123-20

Query Match          97.5%; Score 100.4; DB 5; Length 1277;
Best Local Similarity 98.1%; Pred. No. 9e-26;
Matches 101; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1  CGACAGACGGCAGCAGTCCAGCTCTGGTTCTCTCGGTTTATTCTGTTAGAAATGAA 60
Db      1130 CGACAGACGGCAGCAGTCCAGCTCTGGTTCTCTCGGTTTATTCTGTTAGAAATGAA 1189
;
QY      61  ATGTTCCCATTAATAAGGGGCATGAGCCCTTCTCACAATAA 103
Db      1190 ATGTTCCCATTAATAAGGGGCATGAGCCCTTCTCACAATAA 1232
;
RESULT 11
US-10-800-834-20
; Sequence 20, Application US/10800834
; Publication No. US20040146930A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 94 Human Secreted Proteins
; FILE REFERENCE: P2029P1D3
; CURRENT APPLICATION NUMBER: US/10/800,834
; CURRENT FILING DATE: 2004-03-16
; PRIOR APPLICATION NUMBER: 10/115,123
; PRIOR FILING DATE: 2002-04-04
```

```

; PRIOR APPLICATION NUMBER: 09/461,325
; PRIOR FILING DATE: 1999-12-14
; PRIOR APPLICATION NUMBER: PCT/US99/13418
; PRIOR FILING DATE: 1999-06-15
; PRIOR APPLICATION NUMBER: 60/089,507
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089,508
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089,509
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089,510
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/090,112
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090,113
; PRIOR FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 20
; LENGTH: 1277
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1207)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1272)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-800-834-20

Query Match          97.5%; Score 100.4; DB 7; Length 1277;
Best Local Similarity 98.1%; Pred. No. 9e-26;
Matches 101; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGACAGACGGCAGCAGTCCAGCTCTGGTTTCCTTCTCGGTTTATTCGTGTTAGAA 60
|
Db 1130 CGGACAGACGGCAGCAGTCCAGCTCTGGTTTCCTTCTCGGTTTATTCGTGTTAGAA 1189
|
Qy 61 ATGGTTCCTAATAAGGGGCATGAGCCCTTCTCTCAAAA 103
|
Db 1190 ATGGTTCCTAATAAAGGGGCATGAGCCCTTCTCTCAAAA 1232
|

RESULT 12
US-09-860-670-161
; Sequence 161, Application US/09860670
; Patent No. US20020165137A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA127P1
; CURRENT APPLICATION NUMBER: US/09/860,670
; CURRENT FILING DATE: 2001-05-21
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 289
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 161
; LENGTH: 26555
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-860-670-161

Query Match          95.5%; Score 98.4; DB 3; Length 26555;
Best Local Similarity 99.0%; Pred. No. 1.6e-24;
Matches 99; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGACAGACGGCAGCAGTCCAGCTCTGGTTTCCTTCTCGGTTTATTCGTGTTAGAA 60
|
Db 13862 CGGACAGACGGCAGCAGTCCAGCTCTGGTTTCCTTCTCGGTTTATTCGTGTTAGAA 13921
|
Qy 61 ATGGTTCCTAATAAAGGGGCATGAGCCCTTCTCTCAAA 100
|

US-10-101-510-127
; Sequence 127, Application US/10101510
; Publication No. US20030148295A1
; GENERAL INFORMATION:
; APPLICANT: WAN, JACKSON
; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
; FILE REFERENCE: 15117.0012
; CURRENT APPLICATION NUMBER: US/10/101,510
; CURRENT FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/276,947
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 805
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 127
; LENGTH: 136328
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-101-510-127

Query Match          95.5%; Score 98.4; DB 6; Length 136328;
Best Local Similarity 99.0%; Pred. No. 3e-24;
Matches 99; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGACAGACGGCAGCAGTCCAGCTCTGGTTTCCTTCTCGGTTTATTCGTGTTAGAA 60
|
Db 37299 CGGACAGACGGCAGCAGTCCAGCTCTGGTTTCCTTCTCGGTTTATTCGTGTTAGAA 37358
|
Qy 61 ATGGTTCCTAATAAAGGGGCATGAGCCCTTCTCTCAAA 100
|
Db 37359 ATGGTTCCTAATAAAGGGGCATGAGCCCTTCTCTCAAA 37398
|

Db 13922 ATGGTTCCTAATAAAGGGGCATGAGCCCTTCTCTCAAA 13961
|

RESULT 13
US-10-227-646-161
; Sequence 161, Application US/10227646
; Publication No. US20030235829A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA127P1
; CURRENT APPLICATION NUMBER: US/10/227,646
; CURRENT FILING DATE: 2002-08-26
; PRIOR APPLICATION NUMBER: US/09/860,670
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: 60/205,515
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 289
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 161
; LENGTH: 26555
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-227-646-161

Query Match          95.5%; Score 98.4; DB 6; Length 26555;
Best Local Similarity 99.0%; Pred. No. 1.6e-24;
Matches 99; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGACAGACGGCAGCAGTCCAGCTCTGGTTTCCTTCTCGGTTTATTCGTGTTAGAA 60
|
Db 13862 CGGACAGACGGCAGCAGTCCAGCTCTGGTTTCCTTCTCGGTTTATTCGTGTTAGAA 13921
|

Qy 61 ATGGTTCCTAATAAAGGGGCATGAGCCCTTCTCTCAAA 100
|

US-10-101-510-127
; Sequence 127, Application US/10101510
; Publication No. US20030148295A1
; GENERAL INFORMATION:
; APPLICANT: WAN, JACKSON
; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
; FILE REFERENCE: 15117.0012
; CURRENT APPLICATION NUMBER: US/10/101,510
; CURRENT FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/276,947
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 805
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 127
; LENGTH: 136328
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-101-510-127

Query Match          95.5%; Score 98.4; DB 6; Length 136328;
Best Local Similarity 99.0%; Pred. No. 3e-24;
Matches 99; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGACAGACGGCAGCAGTCCAGCTCTGGTTTCCTTCTCGGTTTATTCGTGTTAGAA 60
|
Db 37299 CGGACAGACGGCAGCAGTCCAGCTCTGGTTTCCTTCTCGGTTTATTCGTGTTAGAA 37358
|
Qy 61 ATGGTTCCTAATAAAGGGGCATGAGCCCTTCTCTCAAA 100
|
Db 37359 ATGGTTCCTAATAAAGGGGCATGAGCCCTTCTCTCAAA 37398
|
```

```
RESULT 15
US-10-172-118-1657
; Sequence 1657, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1657
; LENGTH: 2113
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM_016582
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-1657

Query Match      93.6%; Score 96.4; DB 6; Length 2113;
Best Local Similarity 99.0%; Pred. No. 3.1e-24;
Matches 97; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1  CGGACAGCGGCAGCAGTCCAGCTCTGGTTTCCTTCGTTTATTCGTTAGATGAA 60
Db      2016 CGGACAGCGGCAGCAGTCCAGCTCTGGTTTCCTTCGTTTATTCGTTAGATGAA 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      61  ATGGTTCCTCCATAATAAGGGGCATGAGCCCTTCCTCAC 98
Db      2076 ATGGTTCCTCCATAATAAGGGGCATGAGCCCTTCCTCCC 2113
        |||||||||||||||||||||||||||||||||||||||||||||||||
```

Search completed: December 6, 2005, 22:57:10  
Job time : 257.473 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 11:34:11 ; Search time 70.31 Seconds  
(without alignments)  
455.919 Million cell updates/sec

Title: US-09-980-046B-8

Perfect score: 103

Sequence: 1 cggacagagcgcagctcc.....tgagcccttctccacaaaa 103

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3289935 seqs, 155610033 residues

Total number of hits satisfying chosen parameters: 6579870

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA New:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/PC1\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq2:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq3:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	27.8	27.0	1430	6	US-10-750-185-34214
c 2	26.8	26.0	1976	6	US-10-750-185-24699
c 3	26.6	25.8	1743	6	US-10-750-185-47197
c 4	26.4	25.6	1222	6	US-10-750-185-57246
c 5	26.2	25.4	3883	6	US-10-750-185-61102
c 6	25.8	25.0	1573	6	US-10-750-185-45497
c 7	25.8	25.0	1699	6	US-10-750-185-50782
c 8	25.6	24.9	1230	6	US-10-750-185-47021
c 9	25.6	24.9	1230	6	US-10-750-185-41831
c 10	25.4	24.7	2096	6	US-10-750-185-31080
c 11	25.4	24.7	150450	7	US-11-112-908-54
c 12	25.2	24.5	600	6	US-10-750-185-21418
c 13	25.2	24.5	1943	6	US-10-750-185-31066
c 14	24.8	24.1	1599	6	US-10-750-185-33730
c 15	24.8	24.1	2028	6	US-10-750-185-47588
c 16	24.6	23.9	1187	6	US-10-750-185-32572
c 17	24.6	23.9	1597	6	US-10-750-185-33998
c 18	24.6	23.9	173115	7	US-11-112-908-65
c 19	24.4	23.7	748	6	US-10-750-185-51701
c 20	24.4	23.7	919	6	US-10-750-185-45500
c 21	24.2	23.5	3058	6	US-10-750-185-51651
c 22	24	23.3	963	6	US-10-750-185-25505
c 23	24	23.3	1150	6	US-10-750-185-58765

ALIGNMENTS

RESULT 1

US-10-750-185-34214/c  
; Sequence 34214, Application US/10750185  
; Publication No. US20050260603A1  
; GENERAL INFORMATION:  
; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: WM1100-2  
; CURRENT APPLICATION NUMBER: US/10/750,185  
; CURRENT FILING DATE: 2003-12-31  
; PRIOR APPLICATION NUMBER: US 60/437,482  
; PRIOR FILING DATE: 2002-12-31  
; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 34214  
; LENGTH: 1430  
; TYPE: DNA  
; ORGANISM: Bovine 19866881402808  
US-10-750-185-34214

Query Match 27.0%; Score 27.8; DB 6; Length 1430;  
Best Local Similarity 54.4%; Pred. No. 0.79;  
Matches 56; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

Qy 1 CGGACAGCGCAGCAGCTCCGCTGCTTCCTTCCTCGGTTTATCTGTTAGAAATA 60  
Db 188 CAGACAGACCTCAGCAGCCTCAGCACAATAGTAGTGCTCAGGAGGGCTGTCGATGGA 129  
Qy 61 ATGTTTCCCAATAATAGGGGCATGAGCCCTTCCTCACAATAA 103  
Db 128 TGGGGAACACACAGCAGGTATTATCACTCCTCTCTTCCAGAA 86

RESULT 2

US-10-750-185-24699  
; Sequence 24699, Application US/10750185  
; Publication No. US20050260603A1  
; GENERAL INFORMATION:  
; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.

```
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 24699
; LENGTH: 1976
; TYPE: DNA
; ORGANISM: Bovine 19866880757769
US-10-750-185-24699

Query Match      26.0%; Score 26.8; DB 6; Length 1976;
Best Local Similarity 64.5%; Pred. No. 2.1;
Matches 40; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 36 TCTCGGTTTATCTGTAGAAATGAATGGTTCCCATAAATAAGGGCATGAGCCCTTCCT 95
DB 157 TCTCTTGCAATTTGTGAAGATGACAGTCTGACCATTAACAAGAGTATTTCTCTTCCT 216

QY 96 CA 97
DB 217 GA 218

RESULT 3
US-10-750-185-47197
; Sequence 47197, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 47197
; LENGTH: 1743
; TYPE: DNA
; ORGANISM: Bovine 19866880234735
US-10-750-185-47197

Query Match      25.8%; Score 26.6; DB 6; Length 1743;
Best Local Similarity 63.1%; Pred. No. 2.3;
Matches 41; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 22 AGCTCTGGTTTCCTTCGCGTTTATCTGTAGAAATGGTTCCCATAAATAAGGGG 81
DB 1668 AGGCTCGTTTGTGGTGGATTTAATACATTATAATAAATATGTACCTAAATTTGAAG 1727

QY 82 CATGA 86
DB 1728 AATGA 1732

RESULT 4
US-10-750-185-57246
```

```
; Sequence 57246, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 57246
; LENGTH: 1222
; TYPE: DNA
; ORGANISM: Bovine 19866880532291
US-10-750-185-57246

Query Match      25.6%; Score 26.4; DB 6; Length 1222;
Best Local Similarity 59.2%; Pred. No. 2.4;
Matches 45; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 26 CTGCTTTCCTTCCTCGGTTTATCTGTAGAAATGAAATGGTTCCCATAAATAAGGGCATG 85
DB 716 CTGTTCCACCTTGTAGGATTAGTATGAGGTTAACATACACTTGACATAAACATGCTGAATA 775

QY 86 AGCCCTTCCTCAGAAA 101
DB 776 AGCCTTGCTTAATAAA 791

RESULT 5
US-10-750-185-61102
; Sequence 61102, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 61102
; LENGTH: 3883
; TYPE: DNA
; ORGANISM: Bovine 19866881158230
US-10-750-185-61102

Query Match      25.4%; Score 26.2; DB 6; Length 3883;
Best Local Similarity 56.3%; Pred. No. 4.5;
Matches 49; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 16 AGTCCCAGCTCTGGTTTCCTTCGCGTTTATCTGTAGAAATGAAATGGTTCCCATAAAT 75
DB 45 AGTATAAACTCAATTTCTTTCAAGAGTACATCTTTTCCCTGATATCATCTCTTTAACT 104

QY 76 AAGGGGCATGAGCCCTTCCTCAGAAA 102
```

Db 105 CAGGCGCTGGAGGCTTTTAATCAGACAA 131

## RESULT 6

US-10-750-185-45497/c  
; Sequence 45497, Application US/10750185  
; Publication No. US20050260603A1

## ; GENERAL INFORMATION:

; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MM1100-2  
; CURRENT APPLICATION NUMBER: US/10/750,185  
; CURRENT FILING DATE: 2003-12-31  
; PRIOR APPLICATION NUMBER: US 60/437,482  
; PRIOR FILING DATE: 2002-12-31  
; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 45497  
; LENGTH: 1573  
; TYPE: DNA  
; ORGANISM: Bovine 19866880410534

US-10-750-185-45497

Query Match 25.0%; Score 25.8; DB 6; Length 1573;

Best Local Similarity 53.5%; Pred. No. 4.4;

Matches 54; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

Qy 3 GACGACGGCAGCAGCTCCAGCTCGGTTCTCTCGTTTATCTCTAGAAATGAAT 62

Db 179 GACGAGGGCTTAAATAAGTCCCACTTCTCTGTTGATTAATCTGTTGGATTAATC 120

Qy 63 GGTTCCTCATAAATAAGGGGCATGAGCCCTTCTCTCAAAAAA 103

Db 119 TGTAAACCTAATCTGTTGATTAATCATGACCTCCAGAA 79

## RESULT 7

US-10-750-185-50782/c

; Sequence 50782, Application US/10750185

; Publication No. US20050260603A1

## ; GENERAL INFORMATION:

; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MM1100-2  
; CURRENT APPLICATION NUMBER: US/10/750,185  
; CURRENT FILING DATE: 2003-12-31  
; PRIOR APPLICATION NUMBER: US 60/437,482  
; PRIOR FILING DATE: 2002-12-31  
; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 50782  
; LENGTH: 1699  
; TYPE: DNA  
; ORGANISM: Bovine 19866880598678

US-10-750-185-50782

Query Match

Best Local Similarity 25.0%; Score 25.8; DB 6; Length 1699;

Matches 39; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy 43 TTATTCTGTAGAAATGGTTCCATAAATAAGGGGCATGAGCCCTTCTCTCAAAAA 102

Db 1632 TTTTCTGTGAGGTTTAAAGCATACCCCAAAATCAGATGACGTGGACCTTCTCTCGAAA 1573

Qy 103 A 103

Db 1572 A 1572

## RESULT 8

US-10-750-185-47021

; Sequence 47021, Application US/10750185

; Publication No. US20050260603A1

## ; GENERAL INFORMATION:

; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MM1100-2  
; CURRENT APPLICATION NUMBER: US/10/750,185  
; CURRENT FILING DATE: 2003-12-31  
; PRIOR APPLICATION NUMBER: US 60/437,482  
; PRIOR FILING DATE: 2002-12-31  
; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 47021  
; LENGTH: 757  
; TYPE: DNA  
; ORGANISM: Bovine 19866880858356

US-10-750-185-47021

Query Match

Best Local Similarity 24.9%; Score 25.6; DB 6; Length 757;

Matches 37; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 11 GCAGCAGTCCAGCTCTGGTTTCTCTCGGTTTATCTCTGTTAGAAATGAATGGTT 66

Db 562 GCTGAGTCAAGCCTCGGTTTCTCTCATCTGTAAATGGGAGAGAAATGTTT 617

## RESULT 9

US-10-750-185-41831/c

; Sequence 41831, Application US/10750185

; Publication No. US20050260603A1

## ; GENERAL INFORMATION:

; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MM1100-2  
; CURRENT APPLICATION NUMBER: US/10/750,185  
; CURRENT FILING DATE: 2003-12-31  
; PRIOR APPLICATION NUMBER: US 60/437,482  
; PRIOR FILING DATE: 2002-12-31  
; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 41831  
; LENGTH: 1230  
; TYPE: DNA  
; ORGANISM: Bovine 19866881468051

US-10-750-185-41831

Query Match

Best Local Similarity 24.9%; Score 25.6; DB 6; Length 1230;

Matches 46; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

Qy	2	GGACAGACGGACGAGTCCGACCTCTGGTTTCCTCTCGGTTTATTCGTGTAGATGAAA	61
Db	413	GGTGGGAAGATGCGCTTGTGCTACTGTGTCGTGGTTTTTTTATGTCCTTTAGAGTGACT	354
Qy	62	TGTTCCCATAAATAAGGGG	81
Db	353	CCCTCTCATGAATAGGGTG	334

**RESULT 10**

```

US-10-750-185-31080
; Sequence 31080, Application US/10750185
; Publication No. US2005026063A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFIELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFE
; FILE REFERENCE: WM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 31080
; LENGTH: 2096
; TYPE: DNA
; ORGANISM: Bovine 19866881218394
US-10-750-185-31080

```

	Query Match	24.7%	Score 25.4	DB 6	Length 2096
Best Local Similarity	58.7%	Pred. No. 6.9			
Matches 4	Conservative 0	Mismatches 31	Indels 0	Gaps 0	
Qy	2	GGACAGACGGCAGCAGTCCACGCTCTGGTTTCCTCTCGGTTTATCTGTTAGAAATGAAA	61		
Db	137	GTACAGCCGGCAGAAAGCCTCAAAACTGCCTTACTTTTTCTTAAATTAATAAAAAA	196		
Qy	62	TGGTTCCCATAAATA	76		
Db	197	TTTTTTATTGAAGTA	211		

## RESULT 11

```

SEQUENCE 54, Application US/1112908
;
GENERAL INFORMATION:
;
APPLICANT: Harris, Cole
;
APPLICANT: Davis, Lisa M.
;
TITLE OF INVENTION: Breast Cancer Biomarkers
;
FILE REFERENCE: 04-164-US
;
CURRENT APPLICATION NUMBER: US/11/112,908
;
CURRENT FILING DATE: 2005-04-22
;
PRIOR APPLICATION NUMBER: US 60/564,758
;
PRIOR FILING DATE: 2004-04-23
;
PRIOR APPLICATION NUMBER: US 60/575,978
;
PRIOR FILING DATE: 2004-06-01
;
PRIOR APPLICATION NUMBER: US 60/631,702
;
PRIOR FILING DATE: 2004-11-30
;
PRIOR APPLICATION NUMBER: US 60/633,826
;
PRIOR FILING DATE: 2004-12-07
;
NUMBER OF SEQ ID NOS: 511
;
SOFTWARE: PatentIn version 3.3
;
SEQ ID NO 54
;
LENGTH: 150450
;
TYPE: DNA
;
ORGANISM: Homo sapiens

```

US-11-112-908-54

	Query Match	24.7%;	Score 25.4;	DB 7;	Length 150450;
	Best Local Similarity	58.7%;	Pred. No. 35;		
	Matches	44;	Mismatches 0;	Indels 31;	Gaps 0;
	Conservative				
Qy	27	TGTTTTCTCTCGGTTTATCTGTTAGAAATGGTTCCCATAAATAAGGGGCATGA	86		
Db	37550	TGFGATGCTTCACGCTTTGTTCTTTTGGCTTAAAAATGCTCACCTATGAAGGGTCTTTT	37609		
Qy	87	GCCCTTTCCTCACAAA	101		
Db	37610	GTGTTTCCATATAAA	37624		

## RESULT 12

```

US-10-750-185-21418/c
; Sequence 21418, Application US/10750185
; Publication No. US2005026063A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 21418
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Bovine
; MMBT10387
US-10-750-185-21418

```

Query Match	24.5%	Score 25.2;	DB 6;	Length 600;
Best Local Similarity	56.2%;	Pred. No. 5.1;		
Matches	45;	Conservative 1;	Mismatches 34;	Indels 0; Gaps 0;

  

Qy	2	GGACAGCGCAGCAGTCCGAGCTCTGGTTTCCTCTCGGTTTATCTGTTAGAAATGAAA	61
Db	113	GGTTGGAAGGATGCCTTTGTCTGTTCTGTTGTTTTTTTGTAGTCTCTTAGAGTCACT	54
Qy	62	TGGTTCCCAATAAATAGGGG	81
Db	53	CCCTTCTCATGAATGAGRTG	34

## RESULT 13

```

US-10-750-185-31066/c
; Sequence 31066, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DeNise, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM11100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922

```

```
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 31066
; LENGTH: 1943
; TYPE: DNA
; ORGANISM: Bovine 19866880870888
US-10-750-185-31066

Query Match      24.5%; Score 25.2; DB 6; Length 1943;
Best Local Similarity 57.7%; Pred. No. 8;
Matches 45; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

Qy  2  GGACAGAGCGGAGAGTCCAGCTCTGTTTCCTTCTCGGTTTATTCGTTAGAAATGAAA 61
Db  1128 GGACAGAGTCCAGAGTGGGGGACCGCATACCAATTGTTGGAAGCGTCTCTTAGGATGAAA 1069

Qy  62  TGGTTCCTCAATAAATAGG 79
Db  1068 TTGTGTCGTAGGCAGGG 1051

RESULT 14
US-10-750-185-33730/c
; Sequence 33730, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 33730
; LENGTH: 1599
; TYPE: DNA
; ORGANISM: Bovine 19866880783812
US-10-750-185-33730

Query Match      24.1%; Score 24.8; DB 6; Length 1599;
Best Local Similarity 56.0%; Pred. No. 10;
Matches 47; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

Qy  13  AGCAGTCCAGCTCTGGTTTCTTCTCGGTTTATTCGTTAGAAATGTTCCCATTA 72
Db  949  ATCACTCCAGCTCTACAAAATTCATGGCTCCCTCTTTCACGTGGAAGGCTTGCCCAA 890

Qy  73  AATAGGGGCGCATGAGCCCTTCCTC 96
Db  889  AGTGAGAGCCTCAGCTCTGCCCC 866

RESULT 15
US-10-750-185-47588/c
; Sequence 47588, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
```

```
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 47588
; LENGTH: 2028
; TYPE: DNA
; ORGANISM: Bovine 19866881295675
US-10-750-185-47588

Query Match      24.1%; Score 24.8; DB 6; Length 2028;
Best Local Similarity 80.6%; Pred. No. 11;
Matches 29; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy  29  GTTTCCTTCTCGGTTTATTCGTTAGAAATGAAATGG 64
Db  1892 GTTTCATCTGTTATTTGTTCTGTTTAAATGAAATGG 1857

Search completed: December 6, 2005, 23:16:54
Job time : 70.31 secs
```

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 11:47:47 ; Search time 47.1624 Seconds  
(without alignments)  
4239.411 Million cell updates/sec

Title: US-09-980-046B-9

Perfect score: 30

Sequence: 1 cggaaataaaggctgtgtaagacaaaaa 30

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N Geneseq\_21:\*

1: Geneseqn1980s:\*

2: Geneseqn1990s:\*

3: Geneseqn2000s:\*

4: Geneseqn2001as:\*

5: Geneseqn2001bs:\*

6: Geneseqn2002as:\*

7: Geneseqn2002bs:\*

8: Geneseqn2003as:\*

9: Geneseqn2003bs:\*

10: Geneseqn2003cs:\*

11: Geneseqn2003ds:\*

12: Geneseqn2004as:\*

13: Geneseqn2004bs:\*

14: Geneseqn2005s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30	100.0	30	4	AAC89740
2	28.4	94.7	547	4	Aaf75084 Human col
3	28.4	94.7	642	8	Abx63536 Human CDN
4	28.4	94.7	668	2	Aaz77534 Human ova
5	28.4	94.7	670	2	Aaz33612 Human bre
6	28.4	94.7	678	12	Adq17878 Human sof
7	28.4	94.7	1323	6	Abq54427 Human ova
8	28.4	94.7	1389	12	Adq22502 Human sof
9	26.4	88.0	308	13	Adt50717 Cancer re
10	24.4	81.3	509	13	Adul3731 Solid tum
c 11	24	80.0	469	12	Ado55031 Gene #128
12	24	80.0	1036	12	Adi57714 Human bre
c 13	23	76.7	301	3	Adc66000 Human lun
c 14	23	76.7	301	6	Abi49219 Human lun
c 15	23	76.7	301	6	Abq92405 Human lun
c 16	23	76.7	301	9	Ada28394 Human lun
c 17	23	76.7	301	10	Adh36958 Human lun
c 18	23	76.7	301	12	Adm56761 Human lun
c 19	23	76.7	301	14	Adu98453 Lung tumo

c	20	23	76.7	301	14	AEB10261	Aeb10261 Cancer re
c	21	21.2	70.7	535	4	AAL12750	Aal12750 Human bre
c	22	21.2	70.7	905	11	ACN82890	Acn82890 Breast ca
	23	21	70.0	464	4	AAK88758	Aak88758 Human dig
	24	21	70.0	464	5	AAS31792	Aas31792 Human liv
	25	21	70.0	464	6	ABN90147	Abn90147 Human liv
	26	21	70.0	464	11	ADJ14914	Adj14914 Human liv
	27	21	70.0	486	4	AAK91075	Aak91075 Human dig
	28	21	70.0	486	4	AAK91076	Aak91076 Human dig
	29	21	70.0	486	5	AAS32111	Aas32111 Human liv
	30	21	70.0	486	5	AAS32110	Aas32110 Human liv
	31	21	70.0	486	6	ABN90465	Abn90465 Human liv
	32	21	70.0	486	6	ABN90466	Abn90466 Human liv
	33	21	70.0	486	11	ADJ15379	Adj15379 Human liv
	34	21	70.0	486	11	ADJ15378	Adj15378 Human liv
	35	21	70.0	634	6	ABL63301	Abi63301 Breast ca
	36	21	70.0	634	6	ABL62874	Abi62874 Breast ca
	37	21	70.0	634	6	ABL63092	Abi63092 Breast ca
	38	21	70.0	634	6	ABK35563	Abk35563 Gene ISG1
	39	21	70.0	634	6	ABT10900	Abt10900 Human bre
	40	21	70.0	634	6	ABZ35078	Abz35078 Human gen
	41	21	70.0	634	6	ABK83926	Abk83926 Human CDN
	42	21	70.0	634	8	ACC51007	Acc51007 Human bla
	43	21	70.0	634	8	ACC50195	Acc50195 Breast ca
	44	21	70.0	634	8	ACC51210	Acc51210 Human Plk
	45	21	70.0	634	10	ADD19019	Add19019 Human dis

#### ALIGNMENTS

##### RESULT 1

AAC89740

ID AAC89740 standard; cDNA; 30 BP.

AC AAC89740;

XX 12-MAR-2001 (first entry)

XX Human gastrointestinal inflammation-related cDNA, SEQ ID NO: 9.

XX Human; cytostatic; immunomodulator; immunostimulant; vulnerary;

XX anti-inflammatory; neuroprotective; antibacterial; gene therapy;

XX gastrointestinal inflammation; immune system disorder; genetic disorder;

XX cancer; autoimmune disorder; infection; wound healing; ss.

XX Homo sapiens.

XX WO200073324-A2.

XX 07-DEC-2000.

XX 01-JUN-2000; 2000WO-US015191.

XX 01-JUN-1999; 99US-0137058P.

XX (DIGI-) DIGITAL GENE TECHNOLOGIES INC.

XX Youakim A, Dubose RF, Sims JE, Pribyl TM, Hillbush BS, Hasel KW;

XX WPI; 2001-061508/07.

XX New polynucleotides and polypeptides, useful in gene therapy and in diagnosing a pathological condition, e.g. for modulating gene expression disorders.

XX Claim 1; Page 85; 108pp; English.

XX The present sequence is one of a number of isolated human polynucleotides which are useful in gene therapy, and for diagnosing a pathological condition or a susceptibility to it. In particular, the polynucleotides are useful for modulating gene expression in gastrointestinal

CC inflammation. The polynucleotides are useful for chromosome  
 CC identification, controlling gene expression through triple helix  
 CC formation or antisense DNA or RNA, or identifying individuals from minute  
 CC biological samples using DNA-based identification techniques. The  
 CC polynucleotides can also be used as an alternative to restriction  
 CC fragment length polymorphism (RFLP), by determining the actual base-by-  
 CC base DNA sequences of selected portions of an individual's genome. The  
 CC polynucleotides may also be used as molecular weight markers on Southern  
 CC gels, as diagnostic probes for the presence of a specific mRNA, as a  
 CC probe to subtract-out known sequences in the process of discovering novel  
 CC polynucleotides, or as an antigen to elicit an immune response. The  
 CC polypeptides are useful in diagnostic procedures to detect a disorder.  
 CC The polynucleotides and polypeptides are useful for preventing, treating  
 CC or ameliorating immune system disorders, genetic disorders, cancers, some  
 CC autoimmune disorders, or infections. The polynucleotides and polypeptides  
 CC are also useful for differentiating, proliferating or attracting cells,  
 CC leading to the regeneration of tissues, especially in wounds or burns.  
 CC The polypeptides and polynucleotides may also be used as a food additive  
 CC or preservative  
 XX  
 SQ Sequence 30 BP; 15 A; 3 C; 7 G; 5 T; 0 U; 0 Other;

Query Match 100.0%; Score 30; DB 4; Length 30;

Best Local Similarity 100.0%; Pred. No. 0.033;  
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGAATAAAGGCTGTTGTTAAAGACAAAAA 30  
 |||||  
 Db 1 CGGAATAAAGGCTGTTGTTAAAGACAAAAA 30

## RESULT 2

AAF75084  
 ID AAF75084 standard; cDNA; 547 BP.

XX AAF75084;  
 AC  
 XX 10-MAY-2001 (first entry)  
 DT  
 XX Human colon associated protein cDNA sequence #8.

DE Human; colon; cancer; disease; ss.  
 XX  
 XX Homo sapiens.

OS  
 XX WO200112781-A1.  
 PN  
 XX 22-FEB-2001.

XX 11-AUG-2000; 2000WO-US022157.  
 PF  
 XX 13-AUG-1999; 99US-0148680P.

PR  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX PA  
 XX Birse CE, Rosen CA;

PI  
 XX WPI; 2001-147551/15.  
 DR  
 XX

XX Nucleic acids encoding 13 human colon cancer associated polypeptides,  
 PT useful for preventing, diagnosing and/or treating e.g. cancers  
 PT (especially colon cancer), Parkinson's disease and diabetic retinopathy.

PS Claim 1; Page 310-311; 326pp; English.

XX The present invention relates to 13 human colon cancer-associated  
 CC proteins. These proteins and the nucleic acid encoding them may be used  
 CC in the prevention, diagnosis and treatment of diseases associated with  
 CC inappropriate colon cancer-associated protein expression

XX Sequence 547 BP; 139 A; 152 C; 180 G; 76 T; 0 U; 0 Other;

SQ Query Match 94.7%; Score 28.4; DB 4; Length 547;

Best Local Similarity 96.7%; Pred. No. 0.19;  
 Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGAATAAAGGCTGTTGTTAAAGACAAAAA 30  
 |||||  
 Db 488 CGGAATAAAGGCTGTTGTTAAAGACAAAAA 517

## RESULT 3

ABX63536  
 ID ABX63536 standard; cDNA; 642 BP.

XX AC  
 XX ABX63536;

XX 26-FEB-2003 (first entry)  
 DT  
 XX Human cDNA #536 differentially expressed in activated vascular tissue.

XX Human; gene; ss; vascular tissue; cytostatic; atherosclerosis; cardiant;  
 KW hypotensive; antidiabetic; gynaecological; vasotropic; cerebroprotective;  
 KW gene therapy; vascular disease; cancer; coronary; artery disease;  
 KW hypertension; diabetes; pre-eclampsia; restenosis;  
 KW ischaemia-reperfusion injury; stroke.

XX Homo sapiens.

XX US2002137081-A1.

XX 26-SEP-2002.

XX 08-JAN-2002; 2002US-00044090.

XX 28-JUL-2000; 2000US-0222469P.

PR 08-JAN-2001; 2001US-0260483P.

XX (BAND/) BANDMAN O.

XX Bandman O;

XX WPI; 2003-110597/10.

XX Combination for diagnosing, staging, treating, or monitoring the  
 PT progression of treatment of a vascular disease, e.g. atherosclerosis,  
 PT comprises several cDNAs that are differentially expressed in activated  
 PT vascular tissue.

XX Claim 1; Page; 18pp; English.

XX This invention relates to a combination comprising several cDNAs that are  
 CC differentially expressed in activated vascular tissue. The invention also  
 CC discloses a high throughput method for detecting differentially expressed  
 CC cDNAs in a sample. The cDNAs of the invention may have  
 CC antiatherosclerotic; cytostatic; cardiant; hypotensive; antidiabetic;  
 CC gynaecological; vasotropic and cerebroprotective activities and may be  
 CC used in gene therapy. The cDNAs of the invention may be used in a high-  
 CC throughput methods for detecting differential expression of one or more  
 CC cDNAs in a sample, or screening several molecules or compounds to  
 CC identify a molecule or compound that specifically binds a cDNA of the  
 CC invention. A protein encoded by the cDNA may be used to screen several  
 CC molecules or compounds to identify a ligand that specifically binds to  
 CC the protein, or to produce or purify an antibody to the protein that can  
 CC be used to detect a protein in a sample or purify a natural or  
 CC recombinant protein from a sample. The nucleotides may be useful for  
 CC diagnosing, staging, treating, or monitoring the progression of treatment  
 CC of a vascular disease, e.g. atherosclerosis, cancer, coronary artery  
 CC disease, hypertension, diabetes, pre-eclampsia, ischaemia-reperfusion  
 CC injury, restenosis, or stroke. The cDNAs can also be used for large-scale  
 CC genetic or gene expression analysis of several new nucleic acid  
 CC molecules. Antibodies to the proteins encoded by the cDNAs are useful for  
 CC diagnosing pre-pathologic disorders, and chronic or acute diseases  
 CC associated with abnormalities in the expression, amount or distribution  
 CC of the protein. The present sequence represents a cDNA of the invention  
 CC that is differentially expressed in activated vascular tissue. Note: The



CC sequence data for this patent did not form part of the specification, but  
CC was obtained in electronic format directly from USPTO at  
CC <http://seqdata.uspto.gov/sequence.html?DocID=20020137081>

XX SQ Sequence 642 BP; 141 A; 188 C; 219 G; 94 T; 0 U; 0 Other;

Query Match 94.7%; Score 28.4; DB 8; Length 642;  
Best Local Similarity 96.7%; Pred. No. 0.19;  
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGAATAAAGCGCTGTTGTTAAAGACAAAA 30  
Db 606 CGGAATAAAGCGCTGTTGTTAAAGACAAAA 635

## RESULT 4

AAZ77534  
ID AAZ77534 standard; cDNA; 668 BP.

AC AAZ77534;

DT 10-APR-2000 (first entry)

XX Human ovarian tumor cDNA library derived EST fragment 85.

XX Expressed sequence tag; EST; human; ovarian tumor; anticancer;  
KW gene therapy; treatment; ss.

XX Homo sapiens.

XX DE19817557-A1.

XX 21-OCT-1999.

XX 09-APR-1998; 98DE-01017557.

XX 09-APR-1998; 98DE-01017557.

XX (META-) METAGEN GES GENOMFORSCHUNG MBH.

XX Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;

XX WPI; 1999-591920/51.

XX New nucleic acid sequences expressed in ovarian, and some other, cancer  
PT tissues, and derived polypeptides, for treatment of ovarian cancer and  
PT identification of therapeutic agents.

XX Claim 3; Page 212; 310pp; German.

XX This invention describes novel nucleic acid (cDNA) sequences (A) which  
CC have anticancer activity and are highly expressed in ovarian tumor tissue  
CC (and some also in testis and breast cancer tissue). The products of the  
CC invention can be used for gene therapy. (A) are used (i) for recombinant  
CC expression of polypeptides (B) and (ii) to isolate complete genes. (B)  
CC are used (i) to identify agents suitable for treatment of ovarian cancer;  
CC (ii) directly for treating this form of cancer (including expression from  
CC gene therapy vectors) and (iii) for generation of specific antibodies.  
CC (A) are identified by assembling ESTs (expressed sequence tags) from a  
CC particular tissue type before comparison of expression patterns. This  
CC allows a significantly longer fragment of the gene to be revealed, so  
CC should reduce the number of failures associated with the fact that ESTs  
CC from different libraries may represent different parts of the same  
CC unknown gene, distorting the estimated frequency of occurrence in a  
CC particular tissue. AAZ77450-277572 represent the human ovarian tumor cDNA  
CC library derived EST fragments described in the method of the invention  
CC and encode the protein fragments represented in AAY76505-Y76638

XX Sequence 668 BP; 140 A; 197 C; 233 G; 98 T; 0 U; 0 Other;

Query Match 94.7%; Score 28.4; DB 2; Length 668;  
Best Local Similarity 96.7%; Pred. No. 0.19;  
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGAATAAAGCGCTGTTGTTAAAGACAAAA 30  
Db 637 CGGAATAAAGCGCTGTTGTTAAAGACAAAA 666

## RESULT 5

AAZ33612  
ID AAZ33612 standard; cDNA; 670 BP.

AC AAZ33612;

DT 08-DEC-1999 (first entry)

XX Human breast tumour-associated EST 2.

XX Expressed sequence tag; EST; human; breast; cancer; gene therapy;  
KW treatment; tumour; cytostatic; medicament; ss.

XX Homo sapiens.

XX DE19813839-A1.

XX 23-SEP-1999.

XX 20-MAR-1998; 98DE-01013839.

XX 20-MAR-1998; 98DE-01013839.

XX (META-) METAGEN GES GENOMFORSCHUNG MBH.

XX Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E, Rosenthal A;

XX WPI; 1999-528981/45.

XX Human nucleic acid sequences and protein products from tumor breast  
PT tissue, useful for breast cancer therapy.

XX Claim 3; 83; 188pp; German.

XX This invention describes novel human nucleic acid sequences from tumor  
CC breast tissue which have cytostatic activity. The nucleic acid sequences  
CC can be used to produce and isolate full-length gene sequences. They can  
CC be used to express proteins, which can be used as tools to find an  
CC activity against breast cancer. The sequences can be used in sense or  
CC antisense form. They are especially useful for medicaments for gene  
CC therapy to treat breast cancer. AAZ33611-248617 represents expressed  
CC sequence tags described in the method of the invention

XX Sequence 670 BP; 142 A; 197 C; 233 G; 98 T; 0 U; 0 Other;

Query Match 94.7%; Score 28.4; DB 2; Length 670;  
Best Local Similarity 96.7%; Pred. No. 0.19;  
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGAATAAAGCGCTGTTGTTAAAGACAAAA 30  
Db 637 CGGAATAAAGCGCTGTTGTTAAAGACAAAA 666

## RESULT 6

ADQ17878  
ID ADQ17878 standard; DNA; 678 BP.

XX ADQ17878;

DT 26-AUG-2004 (first entry)

XX Human soft tissue sarcoma-upregulated DNA - SEQ ID 695.

XX soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;  
KW ds.

```

OS Homo sapiens.
XX
XX W02004048938-A2.
XX
XX PD 10-JUN-2004.
XX
XX PF 26-NOV-2003; 2003WO-US038193.
XX
XX PR 26-NOV-2002; 2002US-0429739P.
XX
XX PA (PROT-) PROTEIN DESIGN LABS INC.
XX
XX PI Aziz N, Gineburg WM, Zlotnik A;
XX
XX DR WPI; 2004-441208/41.
XX
XX PT Early detection of soft tissue sarcoma comprises determining expression
PT of a gene in a first soft tissue sample and a normal soft tissue sample
PT and comparing the gene expression, also useful in treating soft tissue
PT sarcoma.
XX
XX PS Example 2; SEQ ID NO 695; 210pp; English.
XX
XX CC The invention relates to a novel method for detecting soft tissue sarcoma
CC which comprises obtaining a first soft tissue sample from an individual
CC and a normal soft tissue sample from the same or different individual,
CC determining the expression of a gene in both samples and comparing the
CC expression of the gene in both soft tissue samples, where a higher level
CC of protein expression in the first soft tissue sample indicates the
CC presence of soft tissue sarcoma. The method of the invention has
CC cytotstatic applications and may be useful for detecting soft tissue
CC sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC acid sequences may be useful in diagnostic and screening applications.
CC The current sequence is that of a human soft tissue sarcoma-upregulated
CC DNA of the invention. The current sequence is not shown within the
CC specification per se but was submitted in CD format by the inventor.
XX
XX SQ Sequence 678 BP; 158 A; 189 C; 227 G; 104 T; 0 U; 0 Other;

Query Match 94.7%; Score 28.4; DB 12; Length 678;
Best Local Similarity 96.7%; Pred. No. 0.19;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGAATAAAGGCTGTGTAAAGACAAAA 30
Db 620 CGGAATAAAGGCTGTGTAAAGACAAAA 649

RESULT 7
ABQ54427
ID ABQ54427 standard; cDNA; 1323 BP.
XX
XX AC ABQ54427;
XX
XX DT 22-AUG-2002 (first entry)
XX
XX DE Human ovarian antigen HPAMB04 cDNA, SEQ ID NO:307.
XX
XX KW Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
XX ovarian cancer; breast cancer; tumour; reproductive system disorder;
XX infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
XX PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;
XX inflammatory condition; immune disorder; blood disorder;
XX cardiovascular disorder; respiratory disorder; neurological disorder;
XX gastrointestinal disorder; urinary system disorder; drug screening;
XX gene therapy; chromosome mapping; forensic analysis;
XX antibody preparation; cytostatic; immunomodulatory; neuroprotective;
XX antiinflammatory; gynaecological; reproductive; gene; ss.
XX
XX OS Homo sapiens.
XX
XX PN W0200200677-A1.
XX
XX XX

```

---

```

PD 03-JAN-2002.
XX
XX PF 07-JUN-2001; 2001WO-US018569.
XX
XX PR 07-JUN-2000; 2000US-0209467P.
XX
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX PI Birse CE, Rosen CA;
XX
XX DR WPI; 2002-147878/19.
XX
XX DR P-PSDB; ABP41350.
XX
XX PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,
XX useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian
XX cancer), immune disorders, cardiovascular disorders and neurological
XX diseases.
XX
XX PS Claim 1; SEQ ID NO 307; 2922pp; English.
XX
XX CC The invention relates to 2175 novel human ovarian antigens (ABP41054-
XX ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
XX encompasses polypeptides 90% identical and polynucleotides 95% identical
XX to the sequences of the invention. The invention additionally relates to
XX recombinant vectors and host cells comprising human ovarian antigen
XX polynucleotides, antibodies against human ovarian antigens, and the use
XX of ovarian antigen polynucleotides and polypeptides in diagnosing,
XX treating, prognosing or preventing various ovary and/or breast-related
XX disorders. Such conditions include ovarian cancer and breast cancer, and
XX metastatic tumours of ovarian or breast origin, reproductive system
XX disorders (e.g., infertility, disorders of pregnancy, anovulation,
XX polycystic ovary syndrome, ovarian cysts, and dysmenorrhea), endocrine
XX disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
XX shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
XX vaginitis), immune disorders (e.g., congenital and acquired
XX immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
XX blood-related disorders (e.g., anaemia), cardiovascular disorders,
XX respiratory disorders, neurological disorders, gastrointestinal disorders
XX and urinary system disorders. Ovarian antigen polypeptides and
XX polynucleotides may also be used in screening for compounds which
XX modulate ovarian antigen expression or activity. The polynucleotides may
XX further be used for gene therapy, chromosome mapping, in the
XX identification of individuals and in forensic analysis, and the
XX polypeptides may be used as food additives or to prepare antibodies
XX useful in disease diagnosis, drug targeting and phenotyping. The present
XX sequence represents cDNA encoding a human ovarian antigen of the
XX invention. Note: The sequence data for this patent did not form part of
XX the printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX SQ Sequence 1323 BP; 343 A; 349 C; 394 G; 235 T; 0 U; 2 Other;

Query Match 94.7%; Score 28.4; DB 6; Length 1323;
Best Local Similarity 96.7%; Pred. No. 0.2;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGAATAAAGGCTGTGTAAAGACAAAA 30
Db 659 CGGAATAAAGGCTGTGTAAAGACAAAA 688

RESULT 8
ADQ22502
ID ADQ22502 standard; DNA; 1389 BP.
XX
XX AC ADQ22502;
XX
XX DT 26-AUG-2004 (first entry)
XX
XX DE Human soft tissue sarcoma-upregulated DNA - SEQ ID 5322.
XX
XX KW soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;
XX ds.

```

```
XX OS Homo sapiens.
XX PN WO2004048938-A2.
XX PD 10-JUN-2004.
XX PF 26-NOV-2003; 2003WO-US038193.
XX PR 26-NOV-2002; 2002US-0429739P.
XX PA (PROT-) PROTEIN DESIGN LABS INC.
XX PI Aziz N, Ginsburg WM, Zlotnik A;
XX WPI; 2004-441208/41.
XX PT Early detection of soft tissue sarcoma comprises determining expression
XX of a gene in a first soft tissue sample and a normal soft tissue sample
XX and comparing the gene expression, also useful in treating soft tissue
XX sarcoma.
XX PS Example 2; SEQ ID NO 5322; 210pp; English.
XX CC The invention relates to a novel method for detecting soft tissue sarcoma
XX which comprises obtaining a first soft tissue sample from an individual
XX and a normal soft tissue sample from the same or different individual,
XX determining the expression of a gene in both samples and comparing the
XX expression of the gene in both soft tissue samples, where a higher level
XX of protein expression in the first soft tissue sample indicates the
XX presence of soft tissue sarcoma. The method of the invention has
XX cytosstatic applications and may be useful for detecting soft tissue
XX sarcoma, possibly via gene therapy or vaccine production. The nucleic
XX acid sequences may be useful in diagnostic and screening applications.
XX The current sequence is that of a human soft tissue sarcoma-upregulated
XX DNA of the invention. The current sequence is not shown within the
XX specification per se but was submitted in CD format by the inventor.
XX SQ Sequence 1389 BP; 483 A; 300 C; 367 G; 239 T; 0 U; 0 Other;

Query Match 94.7%; Score 28.4; DB 12; Length 1389;
Best Local Similarity 96.7%; Pred. No. 0.2;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGAATAAAGCGCTGTTGTAAGACAAAA 30
    |||||
Db 554 CGGAATAAAGCGCTGTTGTAAGACAAAA 583

RESULT 9
ADTS0717
ID ADTS0717 standard; DNA; 308 BP.
AC ADTS0717;
DT 13-JAN-2005 (first entry)
XX Cancer related nucleic acid sequence #21.
XX ds; gene; cytostatic; gene therapy; vaccine; diagnosis; breast; colon;
XX lung; ovarian; prostate; cancer.
XX OS Homo sapiens.
XX PN WO2004092338-A2.
XX PD 28-OCT-2004.
XX PF 12-APR-2004; 2004WO-US011104.
XX PR 11-APR-2003; 2003US-0462399P.
XX PR 01-JUL-2003; 2003US-0484333P.
XX

PA (DIAD-) DIADEXUS INC.
XX Macina RA, Turner LR, Sun Y, Tam A;
XX WPI; 2004-766851/75.
XX PT New cancer specific nucleic acid (CaSNA) molecules, useful for
XX diagnosing, monitoring the presence of, or treating a patient with
XX breast, colon, lung, ovarian, or prostate cancer.
XX PS Claim 1; SEQ ID NO 21; 891pp; English.
XX CC The invention relates to an isolated nucleic acid molecule (I)
XX selectively hybridizing to, or comprising at least 95% sequence identity
XX to, any of the 362 nucleotide sequences fully defined in the
XX specification. The nucleic acid molecules and polypeptides are useful for
XX diagnosing, monitoring the presence of, or treating a patient with
XX breast, colon, lung, ovarian, or prostate cancer. This sequence
XX corresponds to a nucleic acid of the invention.
XX SQ Sequence 308 BP; 68 A; 87 C; 110 G; 43 T; 0 U; 0 Other;

Query Match 88.0%; Score 26.4; DB 13; Length 308;
Best Local Similarity 96.4%; Pred. No. 1.1;
Matches 27; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGAATAAAGCGCTGTTGTAAGACAAA 28
    |||||
Db 281 CGGAATAAAGCGCTGTTGTAAGAGAAA 308

RESULT 10
ADUI3731
ID ADUI3731 standard; DNA; 509 BP.
AC ADUI3731;
XX 27-JAN-2005 (first entry)
XX Solid tumour prognosis gene seqid 4170.
XX cytostatic; gene therapy; expression profile; solid tumour;
XX peripheral blood mononuclear cell; PBMC; prognosis; ds.
XX OS Unidentified.
XX PN WO2004097052-A2.
XX PD 11-NOV-2004.
XX PF 29-APR-2004; 2004WO-US013587.
XX PR 29-APR-2003; 2003US-0466067P.
XX PR 23-JAN-2004; 2004US-0538246P.
XX PA (AMHP ) WYETH.
XX PA (STRA/) STRAHS A.
XX Strahs A, Trepicchio WL, Burczynski ME, Twine NC, Slonim DK;
XX Immerman F, Dorner AJ;
XX WPI; 2004-804779/79.
XX A method, useful for prognosing and treating solid tumor, comprises
XX comparing an expression profile of a gene expressed in peripheral blood
XX mononuclear cells to a reference expression profile of a gene.
XX PS Disclosure; Page; 111pp; English.
XX CC The invention describes a method comprising comparing an expression
XX profile of at least one gene in a peripheral blood sample of a patient to
XX at least one reference expression profile of the at least one gene, where
XX the patient has a solid tumour, and each of the gene is differentially
```

CC expressed in peripheral blood mononuclear cells (PBMCs) of a first class  
 CC of patients as compared to PBMCs of a second class of patients, where  
 CC both of the first and second classes of patients have the solid tumour, and  
 CC each of the first and second classes is a subcluster formed by an  
 CC unsupervised clustering analysis of gene expression profiles in PBMCs of  
 CC a population of patients who have the solid tumour, and where the  
 CC majority of the first class of patients has a first clinical outcome, and  
 CC the majority of the second class of patients has a second clinical  
 CC outcome. Also described are: a system comprising (i) a memory or a  
 CC storage medium including data that represent an expression profile of at  
 CC least one gene in a peripheral blood sample of a patient who has a solid  
 CC tumour, (ii) at least another storage medium including data that  
 CC represent at least one reference expression profile of the gene, (iii) a  
 CC program capable of comparing the expression profile to the reference  
 CC expression profile, and (iv) a processor capable of executing the  
 CC program, where expression levels of the gene in peripheral blood  
 CC mononuclear cells of patients who have the solid tumour correlate with  
 CC clinical outcomes of the patients; and a nucleic acid or protein array  
 CC comprising concentrated probes for solid tumour prognosis genes, where  
 CC each of the solid tumour prognosis genes is differentially expressed in  
 CC PBMCs of a first class of patients as compared to PBMCs of a second class  
 CC of patients, where both the first and second classes of patients have a  
 CC solid tumour, and where the first class of patients has a first clinical  
 CC outcome, and the second class of patients has a second clinical outcome.  
 CC The method, system, and array are useful for prognosis and treating  
 CC solid tumours. This sequence represents a solid tumour prognosis gene of  
 CC the invention. Note: the sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic format  
 CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

SQ Sequence 509 BP; 120 A; 95 C; 128 G; 74 T; 0 U; 92 Other;

Query Match 81.3%; Score 24.4; DB 13; Length 509;  
 Best Local Similarity 83.3%; Pred. No. 6.9;  
 Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGAATAAAGGCTGTTGTAAGACAAAA 30  
 |||||  
 Db 236 CGGAATAAAGGCTGTTGTAANNNGAAAA 265

RESULT 11

AD055031/c

ID AD055031 standard; DNA; 469 BP.

XX AC AD055031;

XX DT 15-JUL-2004 (first entry)

XX DE Gene #128 with increased gene expression in renal cell carcinoma.

XX KW ds; gene; cytostatic; gene therapy; differential expression;

XX KW renal cell carcinoma; clear cell RCC; papillary RCC;

XX KW chromophobe/oncocytoma RCC; sarcomatoid RCC; TCC; Wilms' tumor;

XX KW gene expression; kidney cancer; diagnostic marker; cancer.

XX OS Homo sapiens.

XX PN WO2004032842-A2.

XX PD 22-APR-2004.

XX PF 06-OCT-2003; 2003WO-US031476.

XX PR 04-OCT-2002; 2002US-0415775P.

XX PA (VAND-) VAN ANDEL INST.

XX PI Teh BT, Takahashi M;

XX XX WPI; 2004-340789/31.

XX PT New nucleic acid and polypeptide compositions, useful in the field of

PT molecular biology and medicine, in particular for gene expression  
 PT profiling, identifying diagnostic markers, and treating certain types of  
 PT kidney cancer.

XX Example IV; SEQ ID NO 128; 53pp; English.

XX The invention relates to novel genes that are differentially expressed in  
 CC sub-types of renal cell carcinomas and methods of detecting them using  
 CC nucleic acids and probes. The nucleic acid probes hybridize with part or  
 CC all of a coding sequence that is overexpressed in clear cell renal cell  
 CC carcinoma (CC-RCC), papillary RCC, chromophobe/oncocytoma RCC,  
 CC sarcomatoid RCC, TCC, or Wilms' tumors, which overexpression is based on  
 CC comparison to a baseline value. The methods and compositions of the  
 CC present invention are useful in the field of molecular biology and  
 CC medicine, in particular for gene expression profiling of certain types of  
 CC kidney cancer, in identifying diagnostic markers, and treating such  
 CC cancer patients. This sequence corresponds to a gene with increased  
 CC expression in CC-RCC.

SQ Sequence 469 BP; 69 A; 163 C; 139 G; 98 T; 0 U; 0 Other;

Query Match 80.0%; Score 24; DB 12; Length 469;  
 Best Local Similarity 100.0%; Pred. No. 9.9;  
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGAATAAAGGCTGTTGTAAGA 24

Db 25 CGGAATAAAGGCTGTTGTAAGA 2

RESULT 12

AD157714

ID AD157714 standard; cDNA; 1036 BP.

XX AC AD157714;

XX DT 22-APR-2004 (first entry)

XX DE Human breast specific nucleic acid (BSNA) #85.

XX KW Human; breast specific nucleic acid; BSNA; gene; ss; metastasis;

XX KW breast cancer; cytostatic.

XX OS Homo sapiens.

XX PN WO2003106648-A2.

XX PD 24-DEC-2003.

XX PF 16-JUN-2003; 2003WO-US018934.

XX PR 14-JUN-2002; 2002US-0389327P.

XX PA (DIAD-) DIADEXUS INC.

XX PI Salceda S, Macina RA, Turner LR, Sun Y, Liu C;

XX DR WPI; 2004-082185/08.

XX Novel isolated polypeptide comprising breast specific protein sequences,  
 PT useful for diagnosing or monitoring presence and metastases of breast  
 PT cancer in patient.

XX PS Claim 1; SEQ ID NO 85; 370pp; English.

XX The invention relates to human breast specific nucleic acids (BSNA) and  
 CC the breast specific proteins (BSP) they encode. The nucleic acids are  
 CC useful for determining the presence of a BSNA in a sample which involves  
 CC contacting the sample with a BSNA under conditions in which the BSNA will  
 CC selectively hybridise to a BSNA in the sample, and detecting the  
 CC hybridisation. The nucleic acids are useful for determining the presence  
 CC of a BSP in a sample which involves contacting the sample with suitable  
 CC reagent under conditions in which the reagent will selectively interact

CC with the BSP, and detecting the interaction of the reagent with a BSP in  
CC the sample. The nucleic acids and proteins are useful for diagnosing or  
CC monitoring the presence and metastases of breast cancer in a patient,  
CC which involves determining an amount of nucleic acid or protein and  
CC comparing the determined amount of nucleic acid or protein in the sample  
CC of the patient to the amount of a breast specific marker in a normal  
CC control, where a difference in the determined amount in the sample  
CC compared to the amount in the control is associated with the presence of  
CC breast cancer. The sequences are useful for treating a patient with  
CC breast cancer, involving administering a composition consisting of a BSNA  
CC or a BSP to a patient, where the administration induces an immune  
CC response against the breast cancer cell expressing the BSNA or BSP. This  
CC sequence represents a human BSNA of the invention.

XX SQ Sequence 1036 BP; 199 A; 298 C; 369 G; 170 T; 0 U; 0 Other;

Query Match 80.0%; Score 24; DB 12; Length 1036;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGAATAAAGGCTGTTGTAAGA 24  
|||||  
DB 1012 CGGAATAAAGGCTGTTGTAAGA 1035

RESULT 13  
AAC66000/c  
ID AAC66000 standard; cDNA; 301 BP.

XX AC AAC66000;

XX DT 21-FEB-2001 (first entry)

XX DE Human lung cancer-associated cDNA clone 25405.

XX KW Lung cancer; therapy; treatment; human; tumor; immunogenic; cytostatic;  
XX vaccine; detection; ss.

XX OS Homo sapiens.

XX PN WO200061612-A2.

XX PD 19-OCT-2000.

XX PF 03-APR-2000; 2000WO-US008896.

XX PR 02-APR-1999; 99US-00285479.

XX PR 17-DEC-1999; 99US-00466396.

XX PR 30-DEC-1999; 99US-00476496.

XX PR 10-JAN-2000; 2000US-00480884.

XX PR 22-FEB-2000; 2000US-00510376.

XX PA (CORI-) CORIXA CORP.

XX PI Wang T, Fan L;

XX PS WPI; 2000-628399/60.

XX DR Isolated polypeptide comprising an immunogenic portion of a lung tumor  
XX protein is used for detecting and monitoring progression of lung cancer  
XX in a patient.

XX PT Claim 25a; Page 229; 261pp; English.

XX CC This invention describes a novel isolated polypeptide (I) which  
XX comprising an immunogenic portion of a lung tumor protein or variant (P2)  
XX which have cytostatic activity. The polypeptides and polynucleotides are  
XX used in compositions and vaccines to inhibit the development of cancer,  
XX especially lung cancer, in a patient. Methods described in the invention  
XX can be used to monitor the progression of a cancer by carrying out the  
XX detection at subsequent time points and comparing the results from the  
XX different time points. CD4+ and/or CD8+ T-Cells isolated from a patient  
XX are treated with P2, polynucleotides encoding P2 or antigen presenting

CC cells expressing P2 and then administered to the patient to inhibit  
CC development of cancer

XX SQ Sequence 301 BP; 41 A; 109 C; 81 G; 67 T; 0 U; 3 Other;

Query Match 76.7%; Score 23; DB 3; Length 301;  
Best Local Similarity 95.8%; Pred. No. 23;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGAATAAAGGCTGTTGTAAGA 24

DB 25 CGGAATAAAGGCTGTTGTAAGA 2

RESULT 14

ABL49219/c

ID ABL49219 standard; cDNA; 301 BP.

XX AC ABL49219;

XX DT 01-MAY-2002 (first entry)

XX DE Human lung tumour cDNA sequence clone 25405 SEQ ID NO:304.

XX KW Human; lung tumour; lung cancer; cytostatic; immunostimulant; vaccine;  
XX immune response; ss.

XX OS Homo sapiens.

XX PN WO200200174-A2.

XX PD 03-JAN-2002.

XX PF 28-JUN-2001; 2001WO-US021065.

XX PR 28-JUN-2000; 2000US-00606421.

XX PR 02-AUG-2000; 2000US-00630940.

XX PR 21-AUG-2000; 2000US-00643597.

XX PR 15-SEP-2000; 2000US-00662786.

XX PR 09-OCT-2000; 2000US-00685696.

XX PR 12-DEC-2000; 2000US-00735705.

XX PR 07-MAY-2001; 2001US-00850716.

XX PA (CORI-) CORIXA CORP.

XX PI Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;  
XX McNeill PD, Fanger N, Retter MW, Marnerakis M, Fanger GR;  
XX PI Vedvick TS, Carter D, Watanabe Y, Peckham DW;

XX DR WPI; 2002-090513/12.

XX PT Polynucleotides encoding lung tumor polypeptides, useful for treating  
XX lung cancer or stimulating an immune response.

XX PS Example 1; Page 306; 374pp; English.

XX CC The present invention describes human lung tumour proteins. Human lung  
XX tumour proteins and polynucleotides have cytostatic and immunostimulant  
XX activities, and can be used in vaccine production. Compositions  
XX comprising the lung tumour proteins, polynucleotides, antibodies, fusion  
XX proteins, T cell populations, or antigen presenting cells that express  
XX the lung tumour proteins are useful for treating lung cancer or  
XX stimulating an immune response. ABL48959 to ABL49300 and ABB74946 to  
XX ABB75070 represent sequences used in the exemplification of the present  
XX invention

XX SQ Sequence 301 BP; 41 A; 109 C; 81 G; 67 T; 0 U; 3 Other;

Query Match 76.7%; Score 23; DB 6; Length 301;  
Best Local Similarity 95.8%; Pred. No. 23;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGAATAAAGGCTGTTGTAAGA 24



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 10:04:41 ; Search time 13.094 Seconds  
(without alignments)  
4072.612 Million cell updates/sec

Title: US-09-980-046B-9  
Perfect score: 30  
Sequence: 1 cggaaataaaggctgtgtaagacaaaaa 30

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA.\*  
1: /cgn2\_6/ptodata/1/ina/1 COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5 COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/H COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/PCTUS COMB.seq.\*  
7: /cgn2\_6/ptodata/1/ina/PP COMB.seq.\*  
8: /cgn2\_6/ptodata/1/ina/RE COMB.seq.\*  
9: /cgn2\_6/ptodata/1/ina/backfiles.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	23	76.7	301	US-09-643-597-304	Sequence 304, App
C 2	23	76.7	301	US-09-480-884A-304	Sequence 304, App
C 3	23	76.7	301	US-09-542-615A-304	Sequence 304, App
C 4	23	76.7	301	US-09-606-421B-304	Sequence 304, App
C 5	23	76.7	301	US-09-630-940B-304	Sequence 304, App
C 6	23	76.7	301	US-10-007-700-304	Sequence 304, App
C 7	21	70.0	634	US-09-917-254-33	Sequence 33, Appl
C 8	20.6	68.7	245	US-09-270-767-85	Sequence 85, Appl
C 9	20.6	68.7	245	US-09-270-767-15367	Sequence 15367, A
C 10	20.4	68.0	721	US-08-998-416-618	Sequence 618, App
C 11	20.4	68.0	74790	US-09-949-016-15321	Sequence 15321, A
C 12	20	66.7	195	US-09-248-796A-14057	Sequence 14057, A
C 13	20	66.7	273	US-09-107-532A-1668	Sequence 1668, App
C 14	20	66.7	5220	US-08-777-405A-1	Sequence 1, Appli
C 15	20	66.7	5220	US-08-977-871A-1	Sequence 1, Appli
C 16	20	66.7	5220	US-09-225-951-1	Sequence 1, Appli
C 17	20	66.7	5220	US-09-841-341-1	Sequence 1, Appli
C 18	20	66.7	5220	US-10-027-591-1	Sequence 1, Appli
C 19	20	66.7	5220	US-10-337-192-1	Sequence 1, Appli
C 20	20	66.7	5220	US-10-697-912-1	Sequence 1, Appli
C 21	20	66.7	1664976	US-08-916-421B-1	Sequence 1, Appli
C 22	20	66.7	1664976	US-09-692-570-1	Sequence 1, Appli
C 23	19.6	65.3	601	US-09-949-016-186914	Sequence 186914, A
C 24	19.6	65.3	123463	US-09-949-016-17078	Sequence 17078, A

ALIGNMENTS

RESULT 1  
US-09-643-597-304/c  
; Sequence 304, Application US/09643597  
; Patent No. 6426072  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Fan, Liqun  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Bangor, Chaitanya S.  
; APPLICANT: Hosken, Nancy  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Li, Samuel X.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: McNeill, Patricia D.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.455C11  
; CURRENT APPLICATION NUMBER: US/09/643,597  
; CURRENT FILING DATE: 2000-08-21  
; NUMBER OF SEQ ID NOS: 369  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 304  
; LENGTH: 301  
; TYPE: DNA  
; ORGANISM: Homo sapien  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)...(301)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-643-597-304

Query Match 76.7%; Score 23; DB 3; Length 301;  
Best Local Similarity 95.8%; Pred. No. 4.1;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGAATAAAGGCTGTGTTAAAGA 24  
|||  
DB 25 CGNAATAAAGGCTGTGTTAAAGA 2

RESULT 2  
US-09-480-884A-304/c  
; Sequence 304, Application US/09480884A  
; Patent No. 6482597  
; GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong  
; APPLICANT: Fan, Liqun  
; APPLICANT: Hosken, Nancy A.  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Fanger, Gary R.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY  
; FILE REFERENCE: 210121.455C6  
; CURRENT APPLICATION NUMBER: US/09/480,884A  
; CURRENT FILING DATE: 2001-08-27  
; NUMBER OF SEQ ID NOS: 330  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 304  
; LENGTH: 301  
; TYPE: DNA  
; ORGANISM: Homo sapien  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(301)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-480-884A-304

Query Match 76.7%; Score 23; DB 3; Length 301;  
Best Local Similarity 95.8%; Pred. No. 4.1;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGAATAAAGGCTGTTGTAAGA 24  
Db 25 CGNAATAAAGGCTGTTGTAAGA 2

RESULT 3  
US-09-542-615A-304/c  
; Sequence 304, Application US/09542615A  
; Patent No. 6518256  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Fan, Liqun  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Hosken, Nancy A.  
; APPLICANT: Fanger, Gary R.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY  
; FILE REFERENCE: 210121.455C8  
; CURRENT APPLICATION NUMBER: US/09/542,615A  
; CURRENT FILING DATE: 2000-04-14  
; NUMBER OF SEQ ID NOS: 350  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 304  
; LENGTH: 301  
; TYPE: DNA  
; ORGANISM: Homo sapien  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(301)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-542-615A-304

Query Match 76.7%; Score 23; DB 3; Length 301;  
Best Local Similarity 95.8%; Pred. No. 4.1;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGAATAAAGGCTGTTGTAAGA 24  
Db 25 CGNAATAAAGGCTGTTGTAAGA 2

RESULT 4  
US-09-606-421B-304/c  
; Sequence 304, Application US/09606421B  
; Patent No. 6531315  
; GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong  
; APPLICANT: Fan, Liqun  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Hosken, Nancy  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Li, Samuel X.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.455C9  
; CURRENT APPLICATION NUMBER: US/09/606,421B  
; CURRENT FILING DATE: 2000-06-28  
; NUMBER OF SEQ ID NOS: 358  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 304  
; LENGTH: 301  
; TYPE: DNA  
; ORGANISM: Homo sapien  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(301)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-606-421B-304

Query Match 76.7%; Score 23; DB 3; Length 301;  
Best Local Similarity 95.8%; Pred. No. 4.1;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGAATAAAGGCTGTTGTAAGA 24  
Db 25 CGNAATAAAGGCTGTTGTAAGA 2

RESULT 5  
US-09-630-940B-304/c  
; Sequence 304, Application US/09630940B  
; Patent No. 6737514  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Fan, Liqun  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Hosken, Nancy  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Li, Samuel X.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: McNeill, Patricia D.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.455C10  
; CURRENT APPLICATION NUMBER: US/09/630,940B  
; CURRENT FILING DATE: 2000-08-02  
; NUMBER OF SEQ ID NOS: 367  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 304  
; LENGTH: 301  
; TYPE: DNA  
; ORGANISM: Homo sapien  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(301)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-630-940B-304

Query Match 76.7%; Score 23; DB 3; Length 301;  
Best Local Similarity 95.8%; Pred. No. 4.1;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGAATAAAGGCTGTTGTAAGA 24



```
Db 25 CGNAATAAAGGCTGTGTAAAGA 2
|| |||||
Query Match 76.7%; Score 23; DB 3; Length 301;
Best Local Similarity 95.8%; Pred. No. 4.1;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGAATAAAGGCTGTGTAAAGA 24
|| |||||
Db 25 CGNAATAAAGGCTGTGTAAAGA 2

RESULT 7
US-09-917-254-33
; Sequence 33, Application US/09917254
; Patent No. 6703204
; GENERAL INFORMATION:
; APPLICANT: Mutter, George
; APPLICANT: Baak, Jan
; TITLE OF INVENTION: Prognostic Classification of Breast Cancer
; FILE REFERENCE: B0801/7224(JRV)
; CURRENT APPLICATION NUMBER: US/09/917,254
; CURRENT FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: US 60/222,093
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 33
; LENGTH: 634
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-09-917-254-33

Db 25 CGNAATAAAGGCTGTGTAAAGA 2
|| |||||
Query Match 76.7%; Score 23; DB 3; Length 301;
Best Local Similarity 95.8%; Pred. No. 4.1;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGAATAAAGGCTGTGTAAAGA 24
|| |||||
Db 25 CGNAATAAAGGCTGTGTAAAGA 2

RESULT 6
US-10-007-700-304/c
; Sequence 304, Application US/10007700
; Patent No. 6960570
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Li, Samuel X.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Fanger, Neil
; APPLICANT: Retter, Marc W.
; APPLICANT: Durham, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Peckman, David W.
; APPLICANT: Cai, Feng
; APPLICANT: Foy, Teresa M.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C17
; CURRENT APPLICATION NUMBER: US/10/007,700
; CURRENT FILING DATE: 2001-11-30
; NUMBER OF SEQ ID NOS: 469
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 304
; LENGTH: 301
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 23, 104, 192
; OTHER INFORMATION: n = A,T,C or G
US-10-007-700-304

Db 25 CGNAATAAAGGCTGTGTAAAGA 2
|| |||||
Query Match 70.0%; Score 21; DB 3; Length 634;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGAATAAAGGCTGTGTAA 21
|| |||||
Db 608 CGGAATAAAGGCTGTGTAA 628

RESULT 8
US-09-270-767-85/c
; Sequence 85, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 85
; LENGTH: 245
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-85

Db 237 AAATAAAGGCTGGGTACACACAAAA 211
|| |||||
Qy 4 AAATAAAGGCTGTGTAAAGACAAAA 30
|| |||||
Db 237 AAATAAAGGCTGGGTACACACAAAA 211

RESULT 9
US-09-270-767-15367/c
; Sequence 15367, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15367
; LENGTH: 245
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-15367

Db 237 AAATAAAGGCTGGGTACACACAAAA 211
|| |||||
Qy 4 AAATAAAGGCTGTGTAAAGACAAAA 30
|| |||||
Db 237 AAATAAAGGCTGGGTACACACAAAA 211

RESULT 10
US-08-998-416-618
; Sequence 618, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen
```



```
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 1668:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 273 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (B) LOCATION 1...273
; SEQUENCE DESCRIPTION: SEQ ID NO: 1668:
US-09-107-532A-1668

Query Match 66.7%; Score 20; DB 3; Length 273;
Best Local Similarity 82.1%; Pred. No. 66;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 GGAATAAAGGCTGTGTAAAGACAAAA 29
   ||||| ||||| ||||| ||||| |||||
Db 8 GGAAGAAGAACGCTGTGTAAAGACACA 35

RESULT 14
US-08-777-405A-1
; Sequence 1, Application US/08777405A
; Patent No. 5858753
; GENERAL INFORMATION:
; APPLICANT: Chantry, David
; APPLICANT: Hoekstra, Merl F.
; APPLICANT: Holtzman, Douglas A
; TITLE OF INVENTION: No. 5858753el Lipid Kinase
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall O'Toole Gerstein Murray & Borun
; STREET: 6300 Sears Tower/233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/777,405
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5858753and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/33441
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5220 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
```

```
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 196..3327
US-08-777-405A-1

Query Match 66.7%; Score 20; DB 2; Length 5220;
Best Local Similarity 82.1%; Pred. No. 1e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 3 GAAATAAAGGCTGTGTAAAGACAAAA 30
   ||||| ||||| ||||| ||||| |||||
Db 5185 GAAGTAAGGCAGATGAAGAGAAAAAA 5212

RESULT 15
US-08-977-871A-1
; Sequence 1, Application US/08977871A
; Patent No. 5882910
; GENERAL INFORMATION:
; APPLICANT: Chantry, David
; APPLICANT: Hoekstra, Merl F.
; APPLICANT: Holtzman, Douglas A
; TITLE OF INVENTION: No. 5882910el Lipid Kinase
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall O'Toole Gerstein Murray & Borun
; STREET: 6300 Sears Tower/233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/977,871A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/777,405
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5882910and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/33441
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5220 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 196..3327
US-08-977-871A-1

Query Match 66.7%; Score 20; DB 2; Length 5220;
Best Local Similarity 82.1%; Pred. No. 1e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 3 GAAATAAAGGCTGTGTAAAGACAAAA 30
   ||||| ||||| ||||| ||||| |||||
Db 5185 GAAGTAAGGCAGATGAAGAGAAAAAA 5212
```

Search completed: December 6, 2005, 11:47:28

Job time : 15.094 secs

---

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 11:34:11 ; Search time 20.4786 Seconds  
(without alignments)  
455.919 Million cell updates/sec

Title: US-09-980-046B-9  
Perfect score: 30  
Sequence: 1 cggaaataaaggctgttgaagacaaaaa 30

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3289935 seqs, 155610033 residues

Total number of hits satisfying chosen parameters: 6579870

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA New:\*

1: /cgn2\_6/ptodata/1/pubna/US09\_NEW\_PUB.seq.\*  
2: /cgn2\_6/ptodata/1/pubna/US06\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/1/pubna/US07\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubna/US08\_NEW\_PUB.seq.\*  
5: /cgn2\_6/ptodata/1/pubna/PCT\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/1/pubna/US10\_NEW\_PUB.seq.\*  
7: /cgn2\_6/ptodata/1/pubna/US11\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubna/US11\_NEW\_PUB.seq.\*  
9: /cgn2\_6/ptodata/1/pubna/US11\_NEW\_PUB.seq.\*  
10: /cgn2\_6/ptodata/1/pubna/US60\_NEW\_PUB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	23	76.7	301	6	US-10-623-155-304
C 2	21.2	70.7	940	6	US-10-750-185-47224
C 3	20	66.7	600	6	US-10-750-185-151
C 4	20	66.7	1583	6	US-10-750-185-27698
C 5	20	66.7	2175	6	US-10-750-185-60060
C 6	20	66.7	5220	7	US-11-110-204-1
C 7	19.4	64.7	811	6	US-10-750-185-41266
C 8	19.4	64.7	1490	6	US-10-750-185-42248
C 9	19.4	64.7	1530	6	US-10-750-185-64471
C 10	19.4	64.7	1810	6	US-10-750-185-57881
C 11	19	63.3	19	8	US-11-101-244-663428
C 12	19	63.3	19	8	US-11-101-244-663430
C 13	19	63.3	19	9	US-11-083-784-663428
C 14	19	63.3	19	9	US-11-083-784-663430
C 15	18.8	62.7	1290	6	US-10-467-657-661
C 16	18.8	62.7	1290	6	US-10-467-657-4925
C 17	18.8	62.7	1632	6	US-10-750-185-38307
C 18	18.8	62.7	1956	6	US-10-750-185-45040
C 19	18.8	62.7	4047	6	US-10-750-185-61814
C 20	18.6	62.0	600	6	US-10-750-185-21234
C 21	18.6	62.0	892	6	US-10-750-185-55928
C 22	18.6	62.0	1160	6	US-10-750-185-31088
C 23	18.6	62.0	1305	6	US-10-750-185-51810

C 24	18.4	61.3	1050	6	US-10-793-626-1449	Sequence 1449, Ap
C 25	18.4	61.3	1084	6	US-10-750-185-49927	Sequence 49927, A
C 26	18.4	61.3	1112	6	US-10-750-185-54938	Sequence 54938, A
C 27	18.4	61.3	1113	6	US-10-793-626-1115	Sequence 1115, Ap
C 28	18.4	61.3	1113	6	US-10-793-626-2131	Sequence 2131, Ap
C 29	18.4	61.3	1259	6	US-10-750-185-59326	Sequence 59326, A
C 30	18.4	61.3	1269	6	US-10-793-626-1909	Sequence 1909, Ap
C 31	18.4	61.3	3037	6	US-10-793-626-3904	Sequence 3904, Ap
C 32	18.4	61.3	3180	6	US-10-793-626-3665	Sequence 3665, Ap
C 33	18.4	61.3	3394	6	US-10-793-626-3365	Sequence 3365, Ap
C 34	18.4	61.3	3536	6	US-10-793-626-3957	Sequence 3957, Ap
C 35	18.2	60.7	1987	6	US-10-750-185-25087	Sequence 25087, A
C 36	18	60.0	19	8	US-11-101-244-663437	Sequence 663437, A
C 37	18	60.0	19	9	US-11-083-784-663437	Sequence 663437, A
C 38	18	60.0	1037	6	US-10-750-185-45289	Sequence 45289, A
C 39	18	60.0	1900	6	US-10-750-185-63880	Sequence 63880, A
C 40	18	60.0	1952	6	US-10-750-185-51869	Sequence 51869, A
C 41	18	60.0	2072	6	US-10-750-185-52444	Sequence 52444, A
C 42	18	60.0	2144	6	US-10-750-185-42852	Sequence 42852, A
C 43	18	60.0	2382	6	US-10-750-185-25473	Sequence 25473, A
C 44	18	60.0	150468	7	US-11-112-908-56	Sequence 56, Appl
C 45	18	60.0	193789	7	US-11-112-908-55	Sequence 55, Appl

ALIGNMENTS

RESULT 1  
US-10-623-155-304/c  
; Sequence 304, Application US/10623155  
; Publication No. US20050261166A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tonglong  
; APPLICANT: Peckham, David W.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Fanger, Gary R.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.455C20  
; CURRENT APPLICATION NUMBER: US/10/623.155  
; CURRENT FILING DATE: 2003-07-17  
; NUMBER OF SEQ ID NOS: 560  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 304  
; LENGTH: 301  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 23, 104, 192  
; OTHER INFORMATION: n = A,T,C or G

Query Match 76.7%; Score 23; DB 6; Length 301;  
Best Local Similarity 95.8%; Pred. No. 1.1;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGAATAAAGCGCTGTGTAAGA 24  
|||  
DB 25 CGNAATAAAGCGCTGTGTAAGA 2

RESULT 2  
US-10-750-185-47224/c  
; Sequence 47224, Application US/10750185  
; Publication No. US20050260603A1  
; GENERAL INFORMATION:  
; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen

```
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; PRIOR FILING DATE: 2003-12-31
; PRIOR FILING DATE: 2003-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 47224
; LENGTH: 940
; TYPE: DNA
; ORGANISM: Bovine 19866881078624
US-10-750-185-47224

Query Match      70.7%; Score 21.2; DB 6; Length 940;
Best Local Similarity 88.5%; Pred. No. 6.6;
Matches 23; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 AAATAAAGGCTGTTGTTAAAGACAAAA 29
Db 323 AACTTAAGGCTGTTGTTGAAGACAAAA 298

RESULT 3
US-10-750-185-151
; Sequence 151, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; PRIOR FILING DATE: 2003-12-31
; PRIOR FILING DATE: 2003-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 151
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Bovine MMBT02951
US-10-750-185-151

Query Match      66.7%; Score 20; DB 6; Length 600;
Best Local Similarity 82.1%; Pred. No. 18;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 GGAATAAAGGCTGTTGTTAAAGACAAAA 29
Db 118 GGTGATAAGGCTATTGAAAGACAAAA 145

RESULT 4
US-10-750-185-27698
; Sequence 27698, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; PRIOR FILING DATE: 2003-12-31
; PRIOR FILING DATE: 2003-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 27698
; LENGTH: 1583
; TYPE: DNA
; ORGANISM: Bovine 19866881268447
US-10-750-185-27698

Query Match      66.7%; Score 20; DB 6; Length 1583;
Best Local Similarity 82.1%; Pred. No. 22;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 GGAATAAAGGCTGTTGTTAAAGACAAAA 29
Db 793 GGTGATAAGGCTATTGAAAGACAAAA 820

RESULT 5
US-10-750-185-60060
; Sequence 60060, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; PRIOR FILING DATE: 2003-12-31
; PRIOR FILING DATE: 2003-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 60060
; LENGTH: 2175
; TYPE: DNA
; ORGANISM: Bovine 19866880609486
US-10-750-185-60060

Query Match      66.7%; Score 20; DB 6; Length 2175;
Best Local Similarity 82.1%; Pred. No. 23;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 GGAATAAAGGCTGTTGTTAAAGACAAAA 30
Db 1428 GAAAAAATAAACTGTTGTTGAAACAAAA 1455

RESULT 6
US-11-110-204-1
; Sequence 1, Application US/11110204
; Publication No. US20050261317A1
; GENERAL INFORMATION:
; APPLICANT: SADHU, Chanchal et al.
; TITLE OF INVENTION: INHIBITORS OF HUMAN PHOSPHATIDYLINOSITOL 3-KINASE DELTA
; FILE REFERENCE: 27866/361708
; CURRENT APPLICATION NUMBER: US/11/110,204
; CURRENT FILING DATE: 2005-04-20
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 5220
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
```



Query Match 64.7%; Score 19.4; DB 6; Length 1810;  
Best Local Similarity 79.3%; Pred. No. 38;  
Matches 23; Conservative 0; Mismatches 0; Gaps 0; Indels 0;

Qy 2 GGAATAAAGGCTGTTCTTAAGACAAAAA 30  
||||| | | | | | | | | | | | | | | |  
Db 1183 GGAATAAAGACTATCTTAAAAA 1211

## RESULT 11

US-11-101-244-663428  
; Sequence 663428, Application US/11101244  
; Publication No. US20050246794A1  
; GENERAL INFORMATION:  
; APPLICANT: Dharmacon, Inc.  
; APPLICANT: Khvorova, Anastasia  
; APPLICANT: Reynolds, Angela  
; APPLICANT: Leake, Devin  
; APPLICANT: Marshall, William  
; APPLICANT: Scaringe, Stephen  
; TITLE OF INVENTION: Functional and Hyperfunctional siRNA  
; FILE REFERENCE: 13499US  
; CURRENT APPLICATION NUMBER: US/11/101.244  
; CURRENT FILING DATE: 2005-04-07  
; PRIOR APPLICATION NUMBER: 60/502,050  
; PRIOR FILING DATE: 2003-09-10  
; PRIOR APPLICATION NUMBER: 60/426,137  
; PRIOR FILING DATE: 2002-11-14  
; NUMBER OF SEQ ID NOS: 1591911  
; SOFTWARE: Proprietary  
; SEQ ID NO 663428  
; LENGTH: 19  
; TYPE: RNA  
; ORGANISM: Homo sapiens  
US-11-101-244-663428

Query Match 63.3%; Score 19; DB 8; Length 19;  
Best Local Similarity 73.7%; Pred. No. 24;  
Matches 14; Conservative 5; Mismatches 0; Gaps 0; Indels 0;

Qy 2 GGAATAAAGGCTGTTGTA 20  
||||| | | | | | | | | | | | | | | |  
Db 1 GGAATAAAGGCTGTTGTA 19

## RESULT 12

US-11-101-244-663430  
; Sequence 663430, Application US/11101244  
; Publication No. US20050246794A1  
; GENERAL INFORMATION:  
; APPLICANT: Dharmacon, Inc.  
; APPLICANT: Khvorova, Anastasia  
; APPLICANT: Reynolds, Angela  
; APPLICANT: Leake, Devin  
; APPLICANT: Marshall, William  
; APPLICANT: Scaringe, Stephen  
; TITLE OF INVENTION: Functional and Hyperfunctional siRNA  
; FILE REFERENCE: 13499US  
; CURRENT APPLICATION NUMBER: US/11/101.244  
; CURRENT FILING DATE: 2005-04-07  
; PRIOR APPLICATION NUMBER: 60/502,050  
; PRIOR FILING DATE: 2003-09-10  
; PRIOR APPLICATION NUMBER: 60/426,137  
; PRIOR FILING DATE: 2002-11-14  
; NUMBER OF SEQ ID NOS: 1591911  
; SOFTWARE: Proprietary  
; SEQ ID NO 663430  
; LENGTH: 19  
; TYPE: RNA  
; ORGANISM: Homo sapiens  
US-11-101-244-663430

Query Match 63.3%; Score 19; DB 8; Length 19;

Best Local Similarity 73.7%; Pred. No. 24;  
Matches 14; Conservative 5; Mismatches 0; Gaps 0; Indels 0;

Qy 3 GAAATAAAGGCTGTTGTAA 21  
||||| | | | | | | | | | | | | | | |  
Db 1 GAAATAAAGGCTGTTGTAA 19

## RESULT 13

US-11-083-784-663428  
; Sequence 663428, Application US/11083784  
; Publication No. US20050245475A1  
; GENERAL INFORMATION:  
; APPLICANT: Dharmacon, Inc.  
; APPLICANT: Khvorova, Anastasia  
; APPLICANT: Reynolds, Angela  
; APPLICANT: Leake, Devin  
; APPLICANT: Marshall, William  
; APPLICANT: Scaringe, Stephen  
; TITLE OF INVENTION: Functional and Hyperfunctional siRNA  
; FILE REFERENCE: 13499US  
; CURRENT APPLICATION NUMBER: US/11/083,784  
; CURRENT FILING DATE: 2005-03-18  
; PRIOR APPLICATION NUMBER: US/10/714,333  
; PRIOR FILING DATE: 2003-11-14  
; PRIOR APPLICATION NUMBER: 60/502,050  
; PRIOR FILING DATE: 2003-09-10  
; PRIOR APPLICATION NUMBER: 60/426,137  
; PRIOR FILING DATE: 2002-11-14  
; NUMBER OF SEQ ID NOS: 1591911  
; SOFTWARE: Proprietary  
; SEQ ID NO 663428  
; LENGTH: 19  
; TYPE: RNA  
; ORGANISM: Homo sapiens  
US-11-083-784-663428

Query Match 63.3%; Score 19; DB 9; Length 19;  
Best Local Similarity 73.7%; Pred. No. 24;  
Matches 14; Conservative 5; Mismatches 0; Gaps 0; Indels 0;

Qy 2 GGAATAAAGGCTGTTGTA 20  
||||| | | | | | | | | | | | | | | |  
Db 1 GGAATAAAGGCTGTTGTA 19

## RESULT 14

US-11-083-784-663430  
; Sequence 663430, Application US/11083784  
; Publication No. US20050245475A1  
; GENERAL INFORMATION:  
; APPLICANT: Dharmacon, Inc.  
; APPLICANT: Khvorova, Anastasia  
; APPLICANT: Reynolds, Angela  
; APPLICANT: Leake, Devin  
; APPLICANT: Marshall, William  
; APPLICANT: Scaringe, Stephen  
; TITLE OF INVENTION: Functional and Hyperfunctional siRNA  
; FILE REFERENCE: 13499US  
; CURRENT APPLICATION NUMBER: US/11/083,784  
; CURRENT FILING DATE: 2005-03-18  
; PRIOR APPLICATION NUMBER: US/10/714,333  
; PRIOR FILING DATE: 2003-11-14  
; PRIOR APPLICATION NUMBER: 60/502,050  
; PRIOR FILING DATE: 2003-09-10  
; PRIOR APPLICATION NUMBER: 60/426,137  
; PRIOR FILING DATE: 2002-11-14  
; NUMBER OF SEQ ID NOS: 1591911  
; SOFTWARE: Proprietary  
; SEQ ID NO 663430  
; LENGTH: 19  
; TYPE: RNA  
; ORGANISM: Homo sapiens



US-11-083-784-663430

Query Match 63.3%; Score 19; DB 9; Length 19;  
Best Local Similarity 73.7%; Pred. No. 24;  
Matches 14; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GAAATAAAGGCTGTGTAA 21  
||||:|||||:|:|  
Db 1 GAAAUAAAGGCGUGUAA 19

RESULT 15

US-10-467-657-661  
; Sequence 661, Application US/10467657  
; Publication No. US20050260581A1  
; GENERAL INFORMATION:  
; APPLICANT: CHIRON Spa  
; APPLICANT: FONTANA Maria Rita  
; APPLICANT: PIZZA Mariagrazia  
; APPLICANT: MASIGNANI Vega  
; APPLICANT: MONACI Elisabetta  
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS  
; FILE REFERENCE:  
; CURRENT APPLICATION NUMBER: US/10/467,657  
; CURRENT FILING DATE: 2003-08-11  
; PRIOR APPLICATION NUMBER: GB-0103424.8  
; PRIOR FILING DATE: 2001-02-12  
; NUMBER OF SEQ ID NOS: 9218  
; SOFTWARE: SeqWin99, version 1.04  
; SEQ ID NO 661  
; LENGTH: 1290  
; TYPE: DNA  
; ORGANISM: Neisseria gonorrhoeae  
US-10-467-657-661

Query Match 62.7%; Score 18.8; DB 6; Length 1290;  
Best Local Similarity 90.9%; Pred. No. 62;  
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GGAATAAAGGCTGTGTAAAG 23  
|||||:|||||:|  
Db 648 GGAATACAGGCTGTCGTAAAG 669

Search completed: December 6, 2005, 23:16:55  
Job time : 21.4786 secs

**This Page Blank (uspto)**

**This Page Blank (uspto)**

**This Page Blank (uspto)**

**This Page Blank (uspto)**

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 11:47:47 ; Search time 133.627 Seconds  
(without alignments)  
4239.411 Million cell updates/sec

Title: US-09-980-046B-10

Perfect score: 85

Sequence: 1 cggtagagtcacatcttcgc.....tgccccactaagtagagaaaa 85

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N\_Geneseq\_21:\*

- 1: geneseqn1980s:\*
- 2: geneseqn1990s:\*
- 3: geneseqn2000s:\*
- 4: geneseqn2001as:\*
- 5: geneseqn2001bs:\*
- 6: geneseqn2002as:\*
- 7: geneseqn2002bs:\*
- 8: geneseqn2003as:\*
- 9: geneseqn2003bs:\*
- 10: geneseqn2003cs:\*
- 11: geneseqn2003ds:\*
- 12: geneseqn2004as:\*
- 13: geneseqn2004bs:\*
- 14: geneseqn2005s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	85	100.0	85	4	AAC89741 Human gas
2	83.4	98.1	451	4	AAS04644 Gene expr
3	83.4	98.1	1126	13	AQ38498 Human SNP
4	83.4	98.1	1143	3	Aaf18060 Lung canc
5	83.4	98.1	1271	10	Adj56466 Human CDN
6	83.4	98.1	1458	13	AQ383779 Human tum
7	83.4	98.1	1539	13	AQ38499 Human SNP
8	83.4	98.1	2027	11	ACN89244 Breast ca
9	83.4	98.1	3674	10	ADJ56452 Human CDN
10	83.4	98.1	66109	6	AB162723 Colon ade
11	83.4	98.1	66109	6	ABN97273 Gene #377
12	83.4	98.1	198285	6	ABK84699 Human CDN
13	83.4	98.1	198285	6	ABN97319 Gene #381
14	83.4	98.1	198285	13	ADR52987 Drug ther
15	81.8	96.2	108	4	AAC89745 Human gas
16	54.6	64.2	11398	6	AB134438 Human imm
17	54.6	64.2	11398	6	AB134439 Human imm
18	28	32.9	1205	10	ADF76527 Novel hum
19	28	32.9	1205	13	ACN38440 Tumour-as

c	20	28	32.9	2379	8	ADA69895	Ada69895 Rice gene
c	21	28	32.9	2379	10	ADC07941	Adc07941 Rice DNA
c	22	27.4	32.2	44770	11	ACN43900	Acn43900 Mouse gen
c	23	27.2	32.0	474	13	ACN47800	Acn47800 Cotton pr
c	24	27	31.8	2489	4	AAI64599	Aai64599 Human tum
c	25	26.8	31.5	1002	2	AAX85552	Aax85552 Probe der
c	26	26.8	31.5	1296	2	AAX85549	Aax85549 CDNA of a
c	27	26.8	31.5	1983	4	ABL21789	Ab121789 Drosophil
c	28	26.8	31.5	4097	4	ABL21788	Ab121788 Drosophil
c	29	26.4	31.1	31449	4	ABL09172	Ab109172 Drosophil
c	30	26	30.6	336	14	ADY21257	Ady21257 Barnase 2
c	31	26	30.6	529	12	ADP64404	Adp64404 GUS (uida
c	32	26	30.6	716	12	ADP64412	Adp64412 Barnase 9
c	33	26	30.6	856	14	ADW95072	Adw95072 SH-EP pro
c	34	26	30.6	857	14	ADW95077	Adw95077 SH-EP pro
c	35	26	30.6	2000	6	ABZ15407	Abz15407 Arabidops
c	36	26	30.6	3768	8	ACA27605	ACA27605 Prokaryot
c	37	25.8	30.4	1390	2	AAX85570	Aax85570 CDNA of a
c	38	25.8	30.4	1458	2	AAQ79936	Aaq79936 Murine Ki
c	39	25.6	30.1	323	3	AAQ26084	Aac26084 Human sec
c	40	25.6	30.1	376	11	ADT96789	Adt96789 Colon can
c	41	25.6	30.1	376	11	ADX43271	Adx43271 Human CDN
c	42	25.6	30.1	637	11	ADT94682	Adt94682 Colon can
c	43	25.6	30.1	637	11	ADX41164	Adx41164 Human CDN
c	44	25.6	30.1	793	8	ABZ52278	Abz52278 Aspergill
c	45	25.6	30.1	3589	10	ADF12420	Adf12420 Human PIG

#### ALIGNMENTS

RESULT 1

AAC89741

ID AAC89741 standard; CDNA; 85 BP.

XX AAC89741;

XX 12-MAR-2001 (first entry)

XX Human gastrointestinal inflammation-related CDNA, SEQ ID NO: 10.

XX Human; cytostatic; immunomodulator; immunostimulant; vulnery;

XX anti-inflammatory; neuroprotective; antibacterial; gene therapy;

XX gastrointestinal inflammation; immune system disorder; genetic disorder;

XX cancer; autoimmune disorder; infection; wound healing; ss.

XX Homo sapiens.

XX WO200073324-A2.

XX 07-DEC-2000.

XX 01-JUN-2000; 2000WO-US015191.

XX 01-JUN-1999; 99US-0137058P.

XX (DIGI-) DIGITAL GENE TECHNOLOGIES INC.

XX Youakim A, Dubose RF, Sims JE, Pribyl TM, Hillbush BS, Hasel KW;

XX WPI; 2001-061508/07.

XX New polynucleotides and polypeptides, useful in gene therapy and in diagnosing a pathological condition, e.g. for modulating gene expression in gastrointestinal inflammation, or for treating cancers or genetic disorders.

XX Claim 1; Page 86; 108pp; English.

XX The present sequence is one of a number of isolated human polynucleotides which are useful in gene therapy, and for diagnosing a pathological condition or a susceptibility to it. In particular, the polynucleotides are useful for modulating gene expression in gastrointestinal

CC inflammation. The polynucleotides are useful for chromosome  
CC identification, controlling gene expression through triple helix  
CC formation or antisense DNA or RNA, or identifying individuals from minute  
CC biological samples using DNA-based identification techniques. The  
CC polynucleotides can also be used as an alternative to restriction  
CC fragment length polymorphism (RFLP), by determining the actual base-by-  
CC base DNA sequences of selected portions of an individual's genome. The  
CC polynucleotides may also be used as molecular weight markers on Southern  
CC gels, as diagnostic probes for the presence of a specific mRNA, as a  
CC probe to subtract-out known sequences in the process of discovering novel  
CC polynucleotides, or as an antigen to elicit an immune response. The  
CC polypeptides are useful in diagnostic procedures to detect a disorder.  
CC The polynucleotides and polypeptides are useful for preventing, treating  
CC or ameliorating immune system disorders, genetic disorders, cancers, some  
CC autoimmune disorders, or infections. The polynucleotides and polypeptides  
CC are also useful for differentiating, proliferating or attracting cells,  
CC leading to the regeneration of tissues, especially in wounds or burns.  
CC The polypeptides and polynucleotides may also be used as a food additive  
CC or preservative  
XX  
SQ Sequence 85 BP; 21 A; 20 C; 20 G; 24 T; 0 U; 0 Other;

Query Match 100.0%; Score 85; DB 4; Length 85;  
Best Local Similarity 100.0%; Pred. NO. 2.2e-21;  
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCCAGGGAAGTTT 60  
DB 1 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCCAGGGAAGTTT 60  
QY 61 CTGGTGCCCCCACTAAGTAGAAAA 85  
DB 61 CTGGTGCCCCCACTAAGTAGAAAA 85

RESULT 2  
AAS04644  
ID AAS04644 standard; cDNA; 451 BP.  
XX  
AC AAS04644;  
XX  
DT 07-SEP-2001 (first entry)  
XX  
DE Gene expression profile sequence #144.  
XX  
KW Gene expression profile; hypersensitivity; DNA microarray;  
KW liver toxicity; hepatitis; tumour formation; immunosuppression;  
KW renal toxicity; glomerulitis; neurotoxicity; leukaemia; dementia;  
KW peripheral neuropathy; hypertension; hypotension; myelosuppression;  
KW retinopathy; inflammation; sensitisation; 88.  
XX  
OS Homo sapiens.  
XX  
PN WO200132928-A2.  
XX  
PD 10-MAY-2001.  
XX  
PF 03-NOV-2000; 2000WO-US030474.  
XX  
PR 05-NOV-1999; 99US-01653398P.  
PR 11-APR-2000; 2000US-0196571P.  
XX  
PA (PHAS-) PHASE-1 MOLECULAR TOXICOLOGY.  
XX  
PI Farr S;  
XX  
XX WPI; 2001-328806/34.  
XX  
XX Identifying hypersensitivity in a subject by obtaining a gene expression  
PT profile of hypersensitivity associated genes and detecting a  
PT predetermined pattern of gene expression of hypersensitivity associated  
PT genes.  
XX

PS Claim 24; Page 176; 222pp; English.  
XX  
CC The sequence represents a cDNA from a gene associated with  
CC hypersensitivity to an agent, the sequence was detected in a sample by use  
CC of a DNA microarray containing genes from a gene expression profile  
CC thought to be associated with hypersensitivity to an agent. The invention  
CC relates to methods of obtaining a gene expression profile of genes  
CC associated with hypersensitivity to an agent involving comparing the gene  
CC expression profile of cells treated with the agent with the gene  
CC expression profile of cells not treated with the agent, and determining  
CC the genes that have altered expression due to exposure to the agent.  
CC Hypersensitivity in a subject can then be detected by comparing the gene  
CC expression profile of the subject with that associated with the  
CC hypersensitivity, usually by hybridisation of a sample of mRNA or cDNA  
CC from the subject to a DNA microarray containing genes from the  
CC hypersensitivity profile. The genes in the profiles are associated with  
CC liver toxicity (e.g. hepatitis), tumour formation, immunosuppression,  
CC renal toxicity (e.g. glomerulitis), neurotoxicity, leukaemia, dementia,  
CC peripheral neuropathy, hyper/hypotension, myelosuppression, retinopathy,  
CC inflammation, and sensitisation  
XX  
SQ Sequence 451 BP; 119 A; 100 C; 135 G; 97 T; 0 U; 0 Other;

Query Match 98.1%; Score 83.4; DB 4; Length 451;  
Best Local Similarity 98.8%; Pred. No. 1.5e-20;  
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCCAGGGAAGTTT 60  
DB 353 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCCAGGGAAGTTT 412  
QY 61 CTGGTGCCCCCACTAAGTAGAAAA 85  
DB 413 CTGGTGCCCCCACTAAGTAGATAA 437  
RESULT 3  
ADQ38498  
ID ADQ38498 standard; DNA; 1126 BP.  
XX  
AC ADQ38498;  
XX  
DT 18-NOV-2004 (first entry)  
XX  
DE Human SNP containing myocardial infarction-associated gene, SEQ ID 161.  
XX  
KW Myocardial infarction; detection; single nucleotide polymorphism; SNP;  
KW cardiant; gene therapy; human; gene; de.  
XX  
OS Homo sapiens.  
XX  
PN WO2004058052-A2.  
XX  
PD 15-JUL-2004.  
XX  
PF 22-DEC-2003; 2003WO-US040978.  
XX  
PR 20-DEC-2002; 2002US-0434778P.  
PR 10-MAR-2003; 2003US-0453135P.  
PR 30-APR-2003; 2003US-0466412P.  
PR 23-SEP-2003; 2003US-0504955P.  
XX  
XX (APPL-) APPLERA CORP.  
PA  
PI Cargill M, Devlin JJ, Takubova O;  
XX  
XX WPI; 2004-533949/51.  
DR P-PSDB; ADQ39326.  
XX  
XX Identifying an individual who has an altered risk for developing  
PT myocardial infarction by detecting a single nucleotide polymorphism in  
PT the individual's nucleic acids.  
XX

PS Claim 7; SEQ ID NO 161; 145pp; English.

XX The invention relates to a novel method for identifying an individual who  
CC has an altered risk for developing myocardial infarction. The method  
CC comprises detecting a single nucleotide polymorphism (SNP) in any one of  
CC the nucleotide sequences given in the specification in the individual's  
CC nucleic acids, where the presence of the SNP is correlated with an  
CC altered risk for myocardial infarction in the individual. The invention  
CC further comprises: an isolated nucleic acid molecule comprising at least  
CC 8 contiguous nucleotides where one of the nucleotides is an SNP given in  
CC the specification or its complement and encoding any one of the amino  
CC acid sequences given in the specification; an isolated polypeptide  
CC comprising an amino acid sequence given in the specification; an antibody  
CC that specifically binds to the polypeptide or its antigen-binding  
CC fragment; an amplified polynucleotide containing an SNP given in the  
CC specification and which is between about 16 and 1000 nucleotides in  
CC length; a kit for detecting an SNP in a nucleic acid, comprising the  
CC polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a  
CC nucleic acid molecule; a method of detecting a variant polypeptide; and a  
CC method for identifying an agent useful in treating or preventing  
CC myocardial infarction. The novel detection method has cardiant activity.  
CC The nucleic acids of the invention may be used in gene therapy. The  
CC method is useful in identifying an individual who has an increased or  
CC decreased risk for developing myocardial infarction and for preparing a  
CC composition for treating or preventing myocardial infarction. This  
CC polynucleotide sequence represents a human myocardial infarction-  
CC associated gene containing one or more SNP's of the invention. Note: This  
CC sequence was not shown in the specification. The sequence has come from  
CC an electronic sequence listing downloaded from the WIPO website.

XX SQ Sequence 1126 BP; 259 A; 279 C; 339 G; 247 T; 0 U; 2 Other;

Query Match 98.1%; Score 83.4; DB 13; Length 1126;

Best Local Similarity 98.8%; Pred. No. 2.1e-20;

Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGTGAAGTGCATCTTCTGCGTGTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 60

Db 1022 CGGTGAAGTGCATCTTCTGCGTGTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 1081

Qy 61 CTGGTGCCCCCACTAAGTAGAAAAA 85

Db 1082 CTGGTGCCCCCACTAAGTAGAATAA 1106

RESULT 4

AAF18060

ID AAF18060 standard; DNA; 1143 BP.

XX

AC AAF18060;

XX 14-MAR-2001 (first entry)

XX Lung cancer associated polynucleotide sequence SEQ ID 79.

DE Human; lung cancer associated protein; neuroprotective; cytostatic;

XX cardioactive; immunomodulatory; muscular active; vulnerary;

XX gastrointestinal; nephrotropic; antiinfective; gynecological;

XX antibacterial; diagnosis; neural disorder; immune disorder; reproductive;

XX proliferative disorder; wound healing; infectious disease; ds.

XX Homo sapiens.

OS WO200055180-A2.

XX 21-SEP-2000.

XX 08-MAR-2000; 2000WO-US005918.

XX 12-MAR-1999; 99US-0124270P.

XX (HUMA-) HUMAN GENOME SCI INC.

PA (ROSE/) ROSEN C A.

XX Ruben SM;

PI WPI; 2000-587514/55.

XX P-PSDB; AAB58184.

XX Lung cancer associated gene sequences, referred to as lung cancer

PT antigens, useful for treatment, prevention, and diagnosis of disorders

XX such as lung cancer.

XX Claim 1; Page 556; 1425pp; English.

XX Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer

XX associated proteins represented in AAB58106 - AAB58548. Lung cancer

XX associated proteins and polynucleotide sequences, their agonists, and

XX antagonists may have neuroprotective; cytostatic; cardioactive;

XX immunomodulatory; muscular active general; vulnerary; gastrointestinal

XX general; nephrotropic; antiinfective; gynecological; or antibacterial

XX activity. The invention also includes antibodies specific for the protein

XX or polynucleotide sequences. The lung cancer associated polynucleotide

XX sequences may be used for detection of lung cancer, chromosome

XX identification, as chromosome markers, and for numerous other diagnostic

XX or research purposes. The proteins may be used to treat disorders such as

XX neural, immune, muscular, reproductive, gastrointestinal, pulmonary, also

XX cardiovascular, renal, and proliferative disorders. The proteins may also

XX be used in the treatment of wounds and infectious diseases.

XX Polynucleotide sequences AAF18425 - AAF18433 and peptide AAB58549 are

XX used in the course of the invention for the identification and

XX characterisation of the polynucleotide and protein sequences

XX SQ Sequence 1143 BP; 276 A; 275 C; 348 G; 242 T; 0 U; 2 Other;

Query Match 98.1%; Score 83.4; DB 3; Length 1143;

Best Local Similarity 98.8%; Pred. No. 2.1e-20;

Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGTGAAGTGCATCTTCTGCGTGTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 60

Db 1006 CGGTGAAGTGCATCTTCTGCGTGTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 1065

Qy 61 CTGGTGCCCCCACTAAGTAGAAAAA 85

Db 1066 CTGGTGCCCCCACTAAGTAGAATAA 1090

RESULT 5

ADJ56466

ID ADJ56466 standard; cDNA; 1271 BP.

XX

AC ADJ56466;

XX 06-MAY-2004 (first entry)

XX Human cDNA differentially expressed in MYCN activated cells SeqID 272.

DE human; differential expression; transactivator; proto-oncogene;

XX neuroblastoma; small cell lung cancer; cytostatic; gene therapy; ss;

XX MYCN activated cell.

XX Homo sapiens.

OS US2003119009-A1.

XX 26-JUN-2003.

XX 25-FEB-2002; 2002US-00084817.

XX 23-FEB-2001; 2001US-0270784P.

XX (STUA/) STUART S G.

XX (NUCH/) NUCHTERN J G.

XX (PLON/) PLON S E.

XX (SHOH/) SHOHET J M.

```

XX PI Stuart SG, Nuchtern JG, Plon SE, Shohet JM;
XX DR WPI; 2003-635698/60.
XX PT New genes regulated by MYCN activation, useful in gene therapy,
XX PT particularly for treating a subject with e.g. neuroblastoma or other
XX PT cancers, or for diagnosing, staging or monitoring the treatment of the
XX PT cancer.
XX PS Claim 1; SEQ ID NO 272; 27pp; English.
XX CC This invention relates to novel isolated cDNAs that are differentially
XX CC expressed in MYCN activated cells. Specifically, it refers to
XX CC polynucleotide sequences that exhibit differential expression patterns in
XX CC cells activated by the transactivator MYCN, where MYCN is a proto-
XX CC oncogene that is amplified in neuroblastoma cells and is common in small
XX CC cell lung cancers. The present invention describes these cDNA molecules
XX CC as useful for in hybridisation assays to detect expression of nucleic
XX CC acids (or complementary nucleic acids) in a present in a given sample, as
XX CC well as for screening assays by identifying molecules or compounds that
XX CC specifically bind the cDNA as a ligand and modulate function or activity.
XX CC Accordingly, these compositions exhibit cytostatic activity and can also
XX CC be used for gene therapy purposes. This polynucleotide sequence is a cDNA
XX CC that is differentially expressed in MYCN activated cells, given in an
XX CC exemplification of the invention. NOTE: This sequence does not appear in
XX CC the printed specification but has been obtained in electronic format from
XX CC the US Patent Office at
XX CC ftp.seqdata.uspto.gov/sequence.html?DocID=20030119009.
XX SQ Sequence 1271 BP; 295 A; 317 C; 379 G; 280 T; 0 U; 0 Other;

Query Match 98.1%; Score 83.4; DB 10; Length 1271;
Best Local Similarity 98.8%; Pred. No. 2.2e-20;
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 60
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 61 CTGGTGCCCCCACTAAGTAGAGAAAA 85
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 1220 CTGGTGCCCCCACTAAGTAGAGAAAA 1244
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 6
ADQ83779
ID ADQ83779 standard; cDNA; 1458 BP.
XX AC ADQ83779;
XX DT 07-OCT-2004 (first entry)
XX DE Human tumour-associated antigenic target (TAT) cDNA sequence #593.
XX KW human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;
XX KW cancer; cell proliferative disorder; gene; ss.
XX OS Homo sapiens.
XX PN W02004060270-A2.
XX PD 22-JUL-2004.
XX PF 15-OCT-2003; 2003WO-US029126.
XX PR 18-OCT-2002; 2002US-0418988P.
XX PA (GETH ) GENENTECH INC.
XX PA (WUTD//) WU T D.
XX PA (ZHOU//) ZHOU Y.
XX PI Wu TD, Zhou Y;

Query Match 98.1%; Score 83.4; DB 13; Length 1458;
Best Local Similarity 98.8%; Pred. No. 2.3e-20;
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 60
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 61 CTGGTGCCCCCACTAAGTAGAGAAAA 85
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 1415 CTGGTGCCCCCACTAAGTAGAGAAAA 1439
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 7
ADQ38499
ID ADQ38499 standard; DNA; 1539 BP.
XX AC ADQ38499;
XX DT 18-NOV-2004 (first entry)
XX DE Human SNP containing myocardial infarction-associated gene, SEQ ID 162.
XX KW Myocardial infarction; detection; single nucleotide polymorphism; SNP;

```

KW cardiant; gene therapy; human; gene; ds.

XX Homo sapiens.

XX WO2004058052-A2.

XX 15-JUL-2004.

XX 22-DEC-2003; 2003WO-US040978.

XX 20-DEC-2002; 2002US-0434778P.

XX 10-MAR-2003; 2003US-0453135P.

XX 30-APR-2003; 2003US-0466412P.

XX 23-SEP-2003; 2003US-0504955P.

XX (APPL-) APPLERA CORP.

XX Cargill M, Devlin JJ, Iakoubova O;

XX WPI; 2004-533949/51.

XX P-PSDB; ABQ39327.

XX Identifying an individual who has an altered risk for developing

XX myocardial infarction by detecting a single nucleotide polymorphism in

XX the individual's nucleic acids.

XX Claim 7; SEQ ID NO 162; 145pp; English.

XX The invention relates to a novel method for identifying an individual who  
CC has an altered risk for developing myocardial infarction. The method  
CC comprises detecting a single nucleotide polymorphism (SNP) in any one of  
CC the nucleotide sequences given in the specification in the individual's  
CC nucleic acids, where the presence of the SNP is correlated with an  
CC altered risk for myocardial infarction in the individual. The invention  
CC further comprises: an isolated nucleic acid molecule comprising at least  
CC 8 contiguous nucleotides where one of the nucleotides is an SNP given in  
CC the specification or its complement and encoding any one of the amino  
CC acid sequences given in the specification; an isolated polypeptide  
CC comprising an amino acid sequence given in the specification; an antibody  
CC that specifically binds to the polypeptide or its antigen-binding  
CC fragment; an amplified polynucleotide containing an SNP given in the  
CC specification and which is between about 16' and 1000 nucleotides in  
CC length; a kit for detecting an SNP in a nucleic acid, comprising the  
CC polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a  
CC nucleic acid molecule; a method of detecting a variant polypeptide; and a  
CC method for identifying an agent useful in treating or preventing  
CC myocardial infarction. The novel detection method has cardiant activity.  
CC The nucleic acids of the invention may be used in gene therapy. The  
CC method is useful in identifying an individual who has an increased or  
CC decreased risk for developing myocardial infarction and for preparing a  
CC composition for treating or preventing myocardial infarction. This  
CC polynucleotide sequence represents a human myocardial infarction-  
CC associated gene containing one or more SNPs of the invention. Note: This  
CC sequence was not shown in the specification. The sequence has come from  
CC an electronic sequence listing downloaded from the WIPO website.

XX Sequence 1539 BP; 346 A; 405 C; 454 G; 332 T; 0 U; 2 Other;

Query Match 98.1%; Score 83.4; DB 13; Length 1539;

Best Local Similarity 98.8%; Pred. No. 2.3e-20;

Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 CGGTGAAGTGCATCTTCGCGTGTCTCTATTGAAACAGCATTTCCCCAGGGAAGTTT 60

Db 1435 CGGTGAAGTGCATCTTCGCGTGTCTCTATTGAAACAGCATTTCCCCAGGGAAGTTT 1494

Oy 61 CTGGGTGCCCCACTAAGTAGAAAA 85

Db 1495 CTGGGTGCCCCACTAAGTAGATAA 1519

RESULT 8

ACN89244

ID ACN89244 standard; DNA; 2027 BP.

XX ACN89244;

XX 02-DEC-2004 (first entry)

XX Breast cancer related marker, seq id 10394.

XX Cancer; breast; tumour; cytostatic; marker; detection; therapy; ds.

XX Homo sapiens.

XX US2003099974-A1.

XX 29-MAY-2003.

XX 18-JUL-2002; 2002US-00198846.

XX 18-JUL-2001; 2001US-0306220P.

XX (MILL-) MILLENNIUM PHARM INC.

XX Lillie J, Xu Y, Wang Y, Steinmann K;

XX WPI; 2003-787014/74.

XX Novel isolated polypeptide associated with breast cancer, useful for  
XX detecting presence of polypeptide in sample, as a marker for breast  
XX cancer.

XX Disclosure; SEQ ID NO 10394; 36pp; English.

XX The invention relates to an isolated polypeptide (I) associated with  
CC breast cancer which is encoded by a nucleic acid molecule comprising a  
CC nucleotide sequence (SI). Further disclosed is an antibody that binds to  
CC the polypeptide of the invention. The activity of the polypeptide of the  
CC invention may be described as cytostatic. The antibody is useful for  
CC detecting the presence of (I) in a sample. Nucleic acid molecules of the  
CC invention are useful in the detection of breast tumours. (I) is useful as  
CC a marker for breast cancer and in breast cancer therapy. Sequences given  
CC in records ACN78851-ACN92934 represent nucleic acid markers associated  
CC with breast cancer. Note: The sequence listing does not form part of the  
CC specification but may be obtained in electronic format from the USPTO web  
CC site at seqdata.uspto.gov/sequence.html?DocID=20030099974

XX Sequence 2027 BP; 553 A; 551 C; 565 G; 354 T; 0 U; 4 Other;

Query Match 98.1%; Score 83.4; DB 11; Length 2027;

Best Local Similarity 98.8%; Pred. No. 2.6e-20;

Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 CGGTGAAGTGCATCTTCGCGTGTCTCTATTGAAACAGCATTTCCCCAGGGAAGTTT 60

Db 1282 CGGTGAAGTGCATCTTCGCGTGTCTCTATTGAAACAGCATTTCCCCAGGGAAGTTT 1341

Oy 61 CTGGGTGCCCCACTAAGTAGAAAA 85

Db 1342 CTGGGTGCCCCACTAAGTAGATAA 1366

RESULT 9

ADJ56452

ID ADJ56452 standard; cDNA; 3674 BP.

XX ADJ56452;

XX 06-MAY-2004 (first entry)

XX Human cDNA differentially expressed in MYCN activated cells SeqID 258.

XX human; differential expression; transactivator; proto-oncogene;

XX neuroblastoma; small cell lung cancer; cytostatic; gene therapy; ss;

XX MYCN activated cell.

```
XX Homo sapiens.
OS
XX
XX US2003119009-A1.
PN
XX
XX 26-JUN-2003.
PD
XX
XX
XX 25-FEB-2002; 2002US-00084817.
PF
XX
XX 23-FEB-2001; 2001US-0270784P.
PR
XX (STUA/) STUART S G.
PA (NUCH/) NUCHTERN J G.
PA (PLON/) PLON S E.
PA (SHOH/) SHOHEIT J M.
XX
XX Stuart SG, Nuchtern JG, Plon SE, Shohet JM;
XX WPI; 2003-635698/60.
XX
XX New genes regulated by MYCN activation, useful in gene therapy,
XX particularly for treating a subject with e.g. neuroblastoma or other
XX cancers, or for diagnosing, staging or monitoring the treatment of the
XX cancer.
XX
XX Claim 1; SEQ ID NO 258; 27pp; English.
XX
XX This invention relates to novel isolated cDNAs that are differentially
XX expressed in MYCN activated cells. Specifically, it refers to
XX polynucleotide sequences that exhibit differential expression patterns in
XX cells activated by the transactivator MYCN, where MYCN is a proto-
XX oncogene that is amplified in neuroblastoma cells and is common in small
XX cell lung cancers. The present invention describes these cDNA molecules
XX as useful for in hybridisation assays to detect expression of nucleic
XX acids (or complementary nucleic acids) in a present in a given sample, as
XX well as for screening assays by identifying molecules or compounds that
XX specifically bind the cDNA as a ligand and modulate function or activity.
XX Accordingly, these compositions exhibit cytostatic activity and can also
XX be used for gene therapy purposes. This polynucleotide sequence is a cDNA
XX that is differentially expressed in MYCN activated cells, given in an
XX exemplification of the invention. NOTE: This sequence does not appear in
XX the printed specification but has been obtained in electronic format from
XX the US Patent Office at
XX ftp.seqdata.uspro.gov/sequence.html?DocID=20030119009.
XX
XX Sequence 3674 BP; 737 A; 1036 C; 1075 G; 826 T; 0 U; 0 Other;
XX
XX Query Match 98.1%; Score 83.4; DB 10; Length 3674;
XX Best Local Similarity 98.8%; Pred. No. 3.1e-20;
XX Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX Qy 1 CGGTGAAGTCATCTTCGCGTGTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 60
XX
XX Db 3501 CGGTGAAGTCATCTTCGCGTGTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 3560
XX
XX
XX Qy 61 CTGGTGCCCCCACTAAGTAGAAAA 85
XX
XX Db 3561 CTGGTGCCCCCACTAAGTAGAATAA 3585
XX
XX
XX RESULT 10
XX ABL62723
XX ID ABL62723 standard; DNA; 66109 BP.
XX
XX AC ABL62723;
XX
XX XX
XX 15-MAY-2002 (first entry)
XX
XX DE Colon adenocarcinoma related gene sequence SEQ ID NO:1060.
XX
XX KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
XX stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
XX cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
XX
```

```
KW gene; ds.
XX
XX Homo sapiens.
XX
XX WO200194629-A2.
PN
XX
XX 13-DEC-2001.
PD
XX
XX 30-MAY-2001; 2001WO-US010838.
PF
XX
XX 05-JUN-2000; 2000US-0209473P.
XX
XX 05-JUN-2000; 2000US-0209531P.
XX
XX 18-SEP-2000; 2000US-0233133P.
XX
XX 18-SEP-2000; 2000US-0233617P.
XX
XX 20-SEP-2000; 2000US-0234009P.
XX
XX 20-SEP-2000; 2000US-0234034P.
XX
XX 22-SEP-2000; 2000US-0234052P.
XX
XX 22-SEP-2000; 2000US-0234509P.
XX
XX 22-SEP-2000; 2000US-0234567P.
XX
XX 25-SEP-2000; 2000US-0234923P.
XX
XX 25-SEP-2000; 2000US-0234924P.
XX
XX 25-SEP-2000; 2000US-0235077P.
XX
XX 25-SEP-2000; 2000US-0235082P.
XX
XX 25-SEP-2000; 2000US-0235134P.
XX
XX 25-SEP-2000; 2000US-0235280P.
XX
XX 26-SEP-2000; 2000US-0235637P.
XX
XX 26-SEP-2000; 2000US-0235638P.
XX
XX 27-SEP-2000; 2000US-0235711P.
XX
XX 27-SEP-2000; 2000US-0235720P.
XX
XX 27-SEP-2000; 2000US-0235840P.
XX
XX 27-SEP-2000; 2000US-0235863P.
XX
XX 28-SEP-2000; 2000US-0236028P.
XX
XX 28-SEP-2000; 2000US-0236032P.
XX
XX 28-SEP-2000; 2000US-0236033P.
XX
XX 28-SEP-2000; 2000US-0236034P.
XX
XX 28-SEP-2000; 2000US-0236109P.
XX
XX 28-SEP-2000; 2000US-0236111P.
XX
XX 29-SEP-2000; 2000US-0236842P.
XX
XX 29-SEP-2000; 2000US-0236891P.
XX
XX 02-OCT-2000; 2000US-0237172P.
XX
XX 02-OCT-2000; 2000US-0237173P.
XX
XX 02-OCT-2000; 2000US-0237278P.
XX
XX 02-OCT-2000; 2000US-0237294P.
XX
XX 02-OCT-2000; 2000US-0237295P.
XX
XX 02-OCT-2000; 2000US-0237316P.
XX
XX 03-OCT-2000; 2000US-0237425P.
XX
XX 03-OCT-2000; 2000US-0237598P.
XX
XX 03-OCT-2000; 2000US-0237604P.
XX
XX 03-OCT-2000; 2000US-0237606P.
XX
XX 03-OCT-2000; 2000US-0237608P.
XX
XX 01-NOV-2000; 2000US-0244867P.
XX
XX 01-NOV-2000; 2000US-0245084P.
XX
XX (AVAL-) AVALON PHARM.
XX
XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
XX Soppet DR, Weaver Z;
XX
XX WPI; 2002-188264/24.
XX
XX Screening for anti-neoplastic agent involves exposing cells to a chemical
XX agent to be tested for anti-neoplastic activity, and determining a change
XX in expression of a gene of a signature gene set.
XX
XX Claim 1; SEQ ID NO 1060; 44pp; English.
XX
XX The present invention describes a method (M1) for screening for an anti-
XX neoplastic agent. The method involves exposing cells to a chemical agent
XX to be tested for anti-neoplastic activity, determining a change in
XX expression of at least one gene (I) of a signature gene set, where (I)
XX comprises a sequence (S) selected from 8447 sequences (given in ABL61664
XX to ABL70110), or is at least 95% identical to (S), where a change in
XX expression is indicative of anti-neoplastic activity. (I) has cytostatic
XX
```



CC activity and can be used in gene therapy. M1 can be used for screening an  
CC anti-neoplastic agent, and can be used for producing a product which is  
CC the data collected with respect to the anti-neoplastic agent as a result  
CC of M1, and the data is sufficient to convey the chemical structure and/or  
CC properties of the agent. M1 can be used in the treatment of cancer such  
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,  
CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell  
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous  
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's  
CC tumour

XX Sequence 66109 BP; 17149 A; 14664 C; 14177 G; 20118 T; 0 U; 1 Other;

Query Match 98.1%; Score 83.4; DB 6; Length 66109;

Best Local Similarity 98.8%; Pred. No. 8.3e-20;

Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGTGAAGTGCATCTTCTGCGTGTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 60

Db 37840 CGGTGAAGTGCATCTTCTGCGTGTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 37899

Qy 61 CTGGTGCCCCCACTAAGTAGAAAA 85

Db 37900 CTGGTGCCCCCACTAAGTAGAATAA 37924

#### RESULT 11

ABN97273  
ID ABN97273 standard; DNA; 66109 BP.

XX AC ABN97273;

XX DT 13-AUG-2002 (first entry)

XX DE Gene #3771 used to diagnose liver cancer.

XX KW Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;  
KW metastatic liver tumour; cytostatic; expression profile; disease state;  
KW disease progression; drug toxicity; drug efficacy; drug metabolism.  
XX OS Homo sapiens.

XX PN WO200229103-A2.

XX PD 11-APR-2002.

XX PF 02-OCT-2001; 2001WO-US030589.

XX PR 02-OCT-2000; 2000US-0237054P.

XX PA (GENE-) GENE LOGIC INC.

XX PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;

XX WPI; 2002-426119/45.

XX DR Diagnosing and detecting the progression of liver cancer, hepatocellular  
XX PT carcinoma or metastatic liver tumor in a patient, involves detecting the  
XX PT level of expression of two or more genes in a liver tissue sample.  
XX PS Claim 1; SEQ ID NO 3771; 298pp; English.

XX CC The invention relates to a novel method for diagnosing and detecting the  
XX CC progression of liver cancer, hepatocellular carcinoma or metastatic liver  
XX CC tumor in a patient, and differentiating metastatic liver cancer from  
XX CC hepatocellular carcinoma in a patient, involving detecting the level of  
XX CC expression of two or more genes represented in ABN93503-ABN97455 in a  
XX CC tissue sample. The method of the invention has hepatotropic, and  
XX CC cytostatic activity. The method is useful for diagnosing and detecting  
XX CC the progression of liver cancer, hepatocellular carcinoma and metastatic  
XX CC liver carcinoma in a patient. The method is useful for identifying  
XX CC expression profiles which serve as useful diagnostic markers as well as  
XX CC markers that can be used to monitor disease states, disease progression,

CC drug toxicity, drug efficacy and drug metabolism. Note: The sequence data  
CC for this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 66109 BP; 17149 A; 14664 C; 14177 G; 20118 T; 0 U; 1 Other;

Query Match 98.1%; Score 83.4; DB 6; Length 66109;

Best Local Similarity 98.8%; Pred. No. 8.3e-20;

Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGTGAAGTGCATCTTCTGCGTGTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 60

Db 37840 CGGTGAAGTGCATCTTCTGCGTGTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 37899

Qy 61 CTGGTGCCCCCACTAAGTAGAAAA 85

Db 37900 CTGGTGCCCCCACTAAGTAGAATAA 37924

#### RESULT 12

ABK84699

ID ABK84699 standard; cDNA; 198285 BP.

XX AC ABK84699;

XX DT 14-AUG-2002 (first entry)

XX DE Human cDNA differentially expressed in granulocytic cells #1270.

XX KW Human; ss; granulocytic cell; DNA chip; bacterial infection;  
KW viral infection; parasitic infection; protozoal infection;  
KW fungal infection; sterile inflammatory disease; psoriasis;  
KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;  
KW cardiac reperfusion injury; renal reperfusion injury; ARDS;  
KW adult respiratory distress syndrome; inflammatory bowel disease;  
KW Crohn's disease; ulcerative colitis; periodontal disease;  
KW granulocyte activation; chronic inflammation; allergy.

XX OS Homo sapiens.

XX PN WO200228999-A2.

XX PD 11-APR-2002.

XX PF 03-OCT-2001; 2001WO-US030821.

XX PR 03-OCT-2000; 2000US-0237189P.

XX PA (GENE-) GENE LOGIC INC.

XX PI Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;

XX WPI; 2002-435328/46.

XX DR Detecting granulocyte activation by detecting differential expression of  
XX PT genes associated with granulocyte activation, which serves as diagnostic  
XX PT markers that is useful for monitoring disease states and drug toxicity.  
XX PS Claim 1; SEQ ID NO 1270; 114pp; English.

XX CC The invention relates to detecting (M1) granulocyte (GC) activation  
XX CC (GCA), by detecting the level of expression of gene(s) (Gs) identified by  
XX CC DNA chip analysis as given in the specification, and comparing the  
XX CC expression level to an expression level in an unactivated GC, where  
XX CC differential expression of Gs is indicative of GCA. Also included are  
XX CC modulating (M2) GA by contacting GC with an agent that alters the  
XX CC expression of at least one gene in Gs; (2) screening (M3) for an agent  
XX CC capable of modulating GCA or an inflammation (especially chronic) in a  
XX CC tissue, an allergic response in a subject, exposure of a subject to a  
XX CC pathogen or sterile inflammatory disease using the gene expression  
XX CC profile; (3) detecting (M4) an inflammation (especially chronic) in a  
XX CC tissue, an allergic response in a subject, exposure of a subject to a

CC pathogen or sterile inflammatory disease, by detecting the level of  
 CC expression in a sample of the tissue of gene(s) from Gs, where the level  
 CC of expression of the gene is indicative of inflammation; (4) treating  
 CC (M5) an inflammation (especially chronic) or in a tissue, an allergic  
 CC response in a subject, exposure of a subject to a pathogen or sterile  
 CC inflammatory disease, by contacting a tissue having inflammation with an  
 CC agent that modulates the expression of gene(s) from Gs in the tissue. M1  
 CC is useful for detecting GCA; M2 is useful for modulating Gs; M3 is useful  
 CC for screening an agent capable of modulating GCA preferably in an  
 CC inflammation in a tissue; M4 is useful for detecting an inflammation  
 CC (especially chronic) in a tissue, an allergic response in a subject,  
 CC exposure of a subject to a pathogen or sterile inflammatory disease (e.g.  
 CC psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thromboeosis,  
 CC cardiac reperfusion injury, renal reperfusion injury, ARDS, adult  
 CC respiratory distress syndrome, inflammatory bowel disease, Crohn's  
 CC disease, ulcerative colitis, periodontal disease; also bacterial  
 CC infection, viral infection, parasitic infection, protozoal infection,  
 CC fungal infection and M5 is useful for treating one of the above  
 CC conditions. The present sequence represents a gene differentially  
 CC expressed in granulocytes. Note: The sequence data for this patent did  
 CC not form part of the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences

SQ Sequence 198285 BP; 55426 A; 41198 C; 41497 G; 60164 T; 0 U; 0 Other;  
 Query Match 98.1%; Score 83.4; DB 6; Length 198285;  
 Best Local Similarity 98.8%; Pred. No. 1.2e-19;  
 Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGTGAAGTGCATCTTCGCGTGTCTCTATTGAAACAGCATTTCCCCCAGGGAAGTTT 60  
 |||||  
 DB 113629 CGGTGAAGTGCATCTTCGCGTGTCTCTATTGAAACAGCATTTCCCCCAGGGAAGTTT 113688

QY 61 CTGGTGCCCCCACTAAGTAGAATAA 85  
 |||||  
 DB 113689 CTGGTGCCCCCACTAAGTAGAATAA 113713

RESULT 13  
 ID ABN97319  
 AC ABN97319 standard; DNA; 198285 BP.  
 AC ABN97319;  
 XX 13-AUG-2002 (first entry)  
 DE Gene #3817 used to diagnose liver cancer.  
 XX  
 XX Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;  
 KW metastatic liver tumour; cytostatic; expression profile; disease state;  
 KW disease progression; drug toxicity; drug efficacy; drug metabolism.  
 XX  
 XX Homo sapiens.  
 XX  
 XX W0200229103-A2.  
 XX  
 XX 11-APR-2002.  
 XX  
 XX 02-OCT-2001; 2001WO-US030589.  
 XX  
 XX 02-OCT-2000; 2000US-0237054P.  
 XX  
 XX (GENE-) GENE LOGIC INC.  
 XX  
 XX Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;  
 XX WPI; 2002-426119/45.  
 XX  
 XX Diagnosing and detecting the progression of liver cancer, hepatocellular  
 PT carcinoma or metastatic liver tumor in a patient, involves detecting the  
 PT level of expression of two or more genes in a liver tissue sample.  
 XX

PS Claim 1; SEQ ID NO 3817; 298pp; English.  
 XX  
 CC The invention relates to a novel method for diagnosing and detecting the  
 CC progression of liver cancer, hepatocellular carcinoma or metastatic liver  
 CC tumour in a patient, and differentiating metastatic liver cancer from  
 CC hepatocellular carcinoma in a patient, involving detecting the level of  
 CC expression of two or more genes represented in ABN93503-ABN97455 in a  
 CC tissue sample. The method of the invention has hepatotropic, and  
 CC cytostatic activity. The method is useful for diagnosing and detecting  
 CC the progression of liver cancer, hepatocellular carcinoma and metastatic  
 CC liver carcinoma in a patient. The method is useful for identifying  
 CC expression profiles which serve as useful diagnostic markers as well as  
 CC markers that can be used to monitor disease states, disease progression,  
 CC drug toxicity, drug efficacy and drug metabolism. Note: The sequence data  
 CC for this patent did not form part of the printed specification, but was  
 CC obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences

SQ Sequence 198285 BP; 55426 A; 41198 C; 41497 G; 60164 T; 0 U; 0 Other;  
 Query Match 98.1%; Score 83.4; DB 6; Length 198285;  
 Best Local Similarity 98.8%; Pred. No. 1.2e-19;  
 Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGTGAAGTGCATCTTCGCGTGTCTCTATTGAAACAGCATTTCCCCCAGGGAAGTTT 60  
 |||||  
 DB 113629 CGGTGAAGTGCATCTTCGCGTGTCTCTATTGAAACAGCATTTCCCCCAGGGAAGTTT 113688

QY 61 CTGGTGCCCCCACTAAGTAGAATAA 85  
 |||||  
 DB 113689 CTGGTGCCCCCACTAAGTAGAATAA 113713

RESULT 14  
 ID ADR52987  
 XX ADR52987 standard; DNA; 198285 BP.  
 AC ADR52987;  
 XX  
 XX 18-NOV-2004 (first entry)  
 DE Drug therapy altered expressed gene #338.  
 XX  
 XX drug activity monitoring; expression profile; gene expression;  
 KW peripheral blood sample; peripheral blood mononuclear cell; drug therapy;  
 KW CCI-779; immunosuppressant; rapamycin; mammalian target of rapamycin;  
 KW mTOR; ds.  
 XX  
 XX Homo sapiens.  
 XX  
 XX W02004072265-A2.  
 XX  
 XX 26-AUG-2004.  
 XX  
 XX 11-FEB-2004; 2004WO-US0004118.  
 XX  
 XX 11-FEB-2003; 2003US-0446133P.  
 PR 03-APR-2003; 2003US-0459782P.  
 PR 23-JAN-2004; 2004US-0538246P.  
 XX  
 XX (AMHP ) WYETH.  
 PA (BURC/) BURCZYNSKI M.  
 PA (TWIN/) TWINE N.  
 PA (DORN/) DORNER A. J.  
 PA (TREP/) TREPICCHIO W L.  
 XX  
 XX Burczynski M, Twine N, Dornier AJ, Trepicchio WL;  
 XX WPI; 2004-642301/62.  
 DR  
 XX Monitoring drug activities in vivo comprises comparing an expression  
 PT profile of a gene in a peripheral blood sample of a patient before and  
 PT after drug therapy.  
 PT

XX Disclosure; SEQ ID NO 338; 136pp; English.

XX The invention relates to a method of monitoring drug activities in vivo

XX by comparing an expression profile of at least one gene in a peripheral

XX blood sample of a patient to a reference expression profile of the at

XX least one gene, where the at least one gene is differentially expressed

XX in peripheral blood mononuclear cells (PBMCs) of patients who have a non-

XX blood disease and are subjected to a drug therapy as compared to PBMCs

XX isolated from the patient before the drug therapy, and where the patient

XX has the non-blood disease and is being treated by the drug therapy. The

XX method, kit, and nucleic acid array are useful for monitoring drug

XX activities in vivo. The drug is especially CCI-779, an ester analogue of

XX the immunosuppressant rapamycin which is a potent inhibitor of the

XX mammalian target of rapamycin (mTOR). This sequence represents a gene

XX expressed in PBMC altered by the drug therapy. (Note: this sequence does

XX no form part of the printed specification but was obtained in electronic

XX format from WIPO at ftp.wipo.int/pub/published\_pct\_sequences/).

XX Sequence 198285 BP; 55426 A; 41198 C; 41497 G; 60164 T; 0 U; 0 Other;

XX

Query Match 98.1%; Score 83.4; DB 13; Length 198285;

Best Local Similarity 98.8%; Pred. No. 1.2e-19;

Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGACACAGCATTTCCCCAGGGAAGTTT 60

Db 113629 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGACACAGCATTTCCCCAGGGAAGTTT 113688

Qy 61 CTGGTGCCCCCACTAAGTAGAATAA 85

Db 113689 CTGGTGCCCCCACTAAGTAGAATAA 113713

RESULT 15

AAAC89745

ID AAC89745 standard; cDNA; 108 BP.

XX AAC89745;

XX

XX 12-MAR-2001 (first entry)

XX

XX Human gastrointestinal inflammation-related cDNA, SEQ ID NO: 14.

XX

XX Human; cytostatic; immunomodulator; immunostimulant; vulnerary;

XX anti-inflammatory; neuroprotective; antibacterial; gene therapy;

XX gastrointestinal inflammation; immune system disorder; genetic disorder;

XX cancer; autoimmune disorder; infection; wound healing; ss.

XX

XX Homo sapiens.

XX

XX WO200073324-A2.

XX

XX 07-DEC-2000.

XX

XX 01-JUN-2000; 2000WO-US015191.

XX

XX 01-JUN-1999; 99US-0137058P.

XX

XX (DIGI-) DIGITAL GENE TECHNOLOGIES INC.

XX

XX Youakim A, Dubose RF, Sims JE, Pribyl TM, Hillbush BS, Hasel KW;

XX

XX "PI; 2001-061508/07.

XX

XX "nucleotides and polypeptides, useful in gene therapy and in

XX a pathological condition, e.g. for modulating gene expression

XX in inflammation, or for treating cancers or genetic

XX

XX English.

XX - number of isolated human polynucleotides

CC which are useful in gene therapy, and for diagnosing a pathological

CC condition or a susceptibility to it. In particular, the polynucleotides

CC are useful for modulating gene expression in gastrointestinal

CC inflammation. The polynucleotides are useful for chromosome

CC identification, controlling gene expression through triple helix

CC formation or antisense DNA or RNA, or identifying individuals from minute

CC biological samples using DNA-based identification techniques. The

CC polynucleotides can also be used as an alternative to restriction

CC fragment length polymorphism (RFLP), by determining the actual base-by-

CC base DNA sequences of selected portions of an individual's genome. The

CC polynucleotides may also be used as molecular weight markers on Southern

CC gels, as diagnostic probes for the presence of a specific mRNA, as a

CC probe to subtract-out known sequences in the process of discovering novel

CC polynucleotides, or as an antigen to elicit an immune response. The

CC polypeptides are useful in diagnostic procedures to detect a disorder.

CC The polynucleotides and polypeptides are useful for preventing, treating

CC or ameliorating immune system disorders, genetic disorders, cancers, some

CC autoimmune disorders, or infections. The polynucleotides and polypeptides

CC are also useful for differentiating, proliferating or attracting cells,

CC leading to the regeneration of tissues, especially in wounds or burns.

CC The polypeptides and polynucleotides may also be used as a food additive

CC or preservative

XX

SQ Sequence 108 BP; 35 A; 20 C; 24 G; 29 T; 0 U; 0 Other;

Query Match 96.2%; Score 81.8; DB 4; Length 108;

Best Local Similarity 97.6%; Pred. No. 3.7e-20;

Matches 83; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGACACAGCATTTCCCCAGGGAAGTTT 60

Db 1 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGACACAGCATTTCCCCAGGGAAGTTT 60

Qy 61 CTGGTGCCCCCACTAAGTAGAATAA 85

Db 61 CTGGTGCCCCCACTAAGTAGAATAA 85

Search completed: December 7, 2005, 00:03:17

Job time : 136.627 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 10:04:41 ; Search time 37.0997 Seconds  
(without alignments)  
4072.612 Million cell updates/sec

Title: US-09-980-046B-10  
Perfect score: 85  
Sequence: 1 cggggaagtgcattcttgc.....tgccccactaagttagaaaaa 85

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA.\*  
1: /cgn2\_6/ptodata/1/ina/1 COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5 COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/H COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/PP COMB.seq.\*  
7: /cgn2\_6/ptodata/1/ina/RE COMB.seq.\*  
8: /cgn2\_6/ptodata/1/ina/RE COMB.seq.\*  
9: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	28.2	33.2	21513	3	US-09-949-016-16695
C 2	28.2	33.2	21513	3	US-09-949-016-16696
C 3	28.2	33.2	21513	3	US-09-949-016-16697
C 4	26.4	31.1	45138	3	US-09-949-016-13027
C 5	26	30.6	856	3	US-09-617-543-3
C 6	26	30.6	10103	3	US-09-949-016-11967
C 7	26	30.6	10103	3	US-09-949-016-13245
C 8	25.6	30.1	323	3	US-09-513-999C-30159
C 9	25.6	30.1	2658	3	US-09-248-796A-324
C 10	25.6	30.1	3601	3	US-09-668-680-2
C 11	25.4	29.9	2223	3	US-09-328-352-3880
C 12	25.4	29.9	818128	3	US-09-949-016-14546
C 13	25.4	29.9	818128	3	US-09-949-016-14547
C 14	25.4	29.9	818128	3	US-09-949-016-14548
C 15	25.4	29.9	818128	3	US-09-949-016-14549
C 16	25.4	29.9	818128	3	US-09-949-016-14550
C 17	25.4	29.9	818128	3	US-09-949-016-14551
C 18	25.4	29.9	818128	3	US-09-949-016-14552
C 19	25.4	29.9	818128	3	US-09-949-016-14553
C 20	25.4	29.9	818128	3	US-09-949-016-14554
C 21	25.4	29.9	818128	3	US-09-949-016-14555
C 22	25.4	29.9	818128	3	US-09-949-016-14556
C 23	25.4	29.9	818128	3	US-09-949-016-14557
C 24	25.4	29.9	818128	3	US-09-949-016-14558

25	25.4	29.9	818128	3	US-09-949-016-14559	Sequence 14559, A
26	25.4	29.9	818128	3	US-09-949-016-14560	Sequence 14560, A
27	25.4	29.9	818128	3	US-09-949-016-14561	Sequence 14561, A
28	25.4	29.9	818128	3	US-09-949-016-14562	Sequence 14562, A
29	25.4	29.9	818128	3	US-09-949-016-14563	Sequence 14563, A
30	25.4	29.9	818128	3	US-09-949-016-14564	Sequence 14564, A
31	25.4	29.9	818128	3	US-09-949-016-14565	Sequence 14565, A
32	25.4	29.9	818128	3	US-09-949-016-14566	Sequence 14566, A
33	25	29.4	5226	3	US-09-484-970B-52	Sequence 52, Appl
C 34	24.8	29.2	601	3	US-09-949-016-43835	Sequence 43835, A
C 35	24.8	29.2	601	3	US-09-949-016-43880	Sequence 43880, A
C 36	24.8	29.2	601	3	US-09-949-016-52452	Sequence 52452, A
C 37	24.8	29.2	601	3	US-09-949-016-52497	Sequence 52497, A
C 38	24.8	29.2	1503	3	US-09-328-352-1525	Sequence 1525, Ap
C 39	24.8	29.2	49164	3	US-09-949-016-12985	Sequence 12985, A
C 40	24.8	29.2	49164	3	US-09-949-016-12986	Sequence 12986, A
C 41	24.8	29.2	49164	3	US-09-949-016-13262	Sequence 13262, A
C 42	24.8	29.2	49164	3	US-09-949-016-13263	Sequence 13263, A
C 43	24.6	28.9	601	3	US-09-949-016-138553	Sequence 138553, A
C 44	24.6	28.9	2253	3	US-09-248-796A-783	Sequence 783, App
45	24.6	28.9	101951	3	US-09-949-016-15648	Sequence 15648, A

ALIGNMENTS

RESULT 1

US-09-949-016-16695/c  
; Sequence 16695, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 16695  
; LENGTH: 21513  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-16695

Query Match 33.2%; Score 28.2; DB 3; Length 21513;  
Best Local Similarity 64.6%; Pred No. 3.1;  
Matches 42; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Qy	9	TGCATCTTCTGCGTGTCTCTATTGAACACAGCATTTCCCGGAGGAGTTCTTGGGTGC	68
Db	7635	TCCATAGCTGCCCTTCTTAAAGCTTGAAGCTCTGAGCTCTGAGTTTTCAGGGTTC	7576
Qy	69	CCAC 73	
Db	7575	CCAC 7571	

RESULT 2

US-09-949-016-16696/c  
; Sequence 16696, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 16696  
; LENGTH: 21513  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-16696

Query Match 33.2%; Score 28.2; DB 3; Length 21513;  
Best Local Similarity 64.6%; Pred. No. 3.1;  
Matches 42; Conservative 0; Mismatches 23; Indels 0; Gaps 0;  
  
QY 9 TGCATCTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTTCTGGGTGC 68  
Db 7635 TCCATAGCCTGCCTTTTCTTTAAATTTCTAAAGCTTTGAGCCTCTGGATTTTTCAGGGTTC 7576  
  
QY 69 CCCAC 73  
Db 7575 CCCAC 7571

RESULT 3  
US-09-949-016-16697/c  
; Sequence 16697, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 16697  
; LENGTH: 21513  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-16697

Query Match 33.2%; Score 28.2; DB 3; Length 21513;  
Best Local Similarity 64.6%; Pred. No. 3.1;  
Matches 42; Conservative 0; Mismatches 23; Indels 0; Gaps 0;  
  
QY 9 TGCATCTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTTCTGGGTGC 68  
Db 7635 TCCATAGCCTGCCTTTTCTTTAAATTTCTAAAGCTTTGAGCCTCTGGATTTTTCAGGGTTC 7576  
  
QY 69 CCCAC 73  
Db 7575 CCCAC 7571

RESULT 4  
US-09-949-016-13027  
; Sequence 13027, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13027  
; LENGTH: 45138  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: (1)...(45138)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-13027

Query Match 31.1%; Score 26.4; DB 3; Length 45138;  
Best Local Similarity 69.2%; Pred. No. 20;  
Matches 36; Conservative 0; Mismatches 16; Indels 0; Gaps 0;  
  
QY 23 GTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTTCTGGGTGCCCACT 74  
Db 15383 GTTATTTATTAAACAAGAATTTCTGACAAAGAAGTTTGTATGTCCTCTCT 15434

RESULT 5  
US-09-617-543-3/c  
; Sequence 3, Application US/09617543  
; Patent No. 6849776  
; GENERAL INFORMATION:  
; APPLICANT: KUVSHINOV, Viktor  
; APPLICANT: KOIVU, Kimmo  
; APPLICANT: KANERVA, Anne  
; APPLICANT: PEHY, Biija  
; TITLE OF INVENTION: MOLECULAR CONTROL OF TRANSGENE SEGREGATION AND ESCAPE  
; BY RECOVERABLE BLOCK OF FUNCTION (RBF) SYSTEM  
; FILE REFERENCE: KUVSHINOV=1  
; CURRENT APPLICATION NUMBER: US/09/617,543  
; CURRENT FILING DATE: 2000-07-14  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 856  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Vigna mungo  
; (SH-EP promoter), Bacillus amyloliquefaciens  
; OTHER INFORMATION: (barnase gene), Escherichia coli (uidA gene)  
; NAME/KEY: promoter  
; LOCATION: (1)..(352)  
; OTHER INFORMATION: Modified SH-EP promoter.  
; NAME/KEY: misc.feature  
; LOCATION: (353)..(404)  
; OTHER INFORMATION: 5' end exon of modified uidA gene.  
; NAME/KEY: misc.feature  
; LOCATION: (443)..(778)  
; OTHER INFORMATION: Synthetic coding sequence of barnase gene.  
; NAME/KEY: intron  
; LOCATION: (405)..(830)  
; OTHER INFORMATION: Intron of uidA gene.  
; NAME/KEY: misc.feature  
; LOCATION: (831)..(856)  
; OTHER INFORMATION: Exon of modified uidA gene.  
US-09-617-543-3

Query Match 30.6%; Score 26; DB 3; Length 856;  
Best Local Similarity 65.5%; Pred. No. 6.5;  
Matches 38; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 24 TTCTCTATTGACACGATTTCCTCCAGGAGTTTCTGGTGCCCACTAAGTAGA 81  
DB 625 TTCTCTATTGAAAAATATCTCTCCATAGATTTCCTGGAGCAACATCAGCAAGA 568

RESULT 6  
US-09-949-016-11967/c  
; Sequence 11967, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11967  
; LENGTH: 10103  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-11967

Query Match 30.6%; Score 26; DB 3; Length 10103;  
Best Local Similarity 59.5%; Pred. No. 16;  
Matches 44; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 7 AGTCATCTCTCGGTGTTCTCTATTGACACGATTTCCTCCAGGAGTTTCTCGGT 66  
DB 8762 AGCTCGGCTCTGGTGCCATGCTGCTTGGAAATGCTCCAGAGCAGGATCTGTGT 8703

QY 67 GCCCCACTAAGTAG 80  
DB 8702 TCCCATTTGAGTTG 8689

RESULT 7  
US-09-949-016-13245/c  
; Sequence 13245, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13245  
; LENGTH: 10103  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-13245

Query Match 30.6%; Score 26; DB 3; Length 10103;

Best Local Similarity 59.5%; Pred. No. 16;  
Matches 44; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 7 AGTCATCTCTCGGTGTTCTCTATTGACACGATTTCCTCCAGGAGTTTCTCGGT 66  
DB 8762 AGCTCGGCTCTGGTGCCATGCTGCTTGGAAATGCTCCAGAGCAGGATCTGTGT 8703

QY 67 GCCCCACTAAGTAG 80  
DB 8702 TCCCATTTGAGTTG 8689

RESULT 8  
US-09-513-999C-30159/c  
; Sequence 30159, Application US/09513999C  
; Patent No. 6783961  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Duclert, A.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
; Patent No. 6783961  
; FILE REFERENCE: 59.US2.REG  
; CURRENT APPLICATION NUMBER: US/09/513,999C  
; CURRENT FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/122,487  
; PRIOR FILING DATE: 1999-02-26  
; NUMBER OF SEQ ID NOS: 36681  
; SOFTWARE: Patent.pm  
; SEQ ID NO 30159  
; LENGTH: 323  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 118  
; OTHER INFORMATION: w=a or t  
US-09-513-999C-30159

Query Match 30.1%; Score 25.6; DB 3; Length 323;  
Best Local Similarity 59.7%; Pred. No. 6.4;  
Matches 43; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 11 CATCTTCTCGGTGTTCTCTATTGACACGATTTCCTCCAGGAGTTTCTGGTGCCC 70  
DB 84 CAGCTTTGCCATATCTTTCCTAATAAGAGCAATCAAGCCAGCAATTTCACAGTTTCC 25

QY 71 CACTAAGTAGAA 82  
DB 24 TACTAAGACAA 13

RESULT 9  
US-09-248-796A-324  
; Sequence 324, Application US/09248796A  
; Patent No. 6747137  
; GENERAL INFORMATION:  
; APPLICANT: Keith Weinstock et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.132  
; CURRENT APPLICATION NUMBER: US/09/248,796A  
; CURRENT FILING DATE: 1999-02-12  
; PRIOR APPLICATION NUMBER: US 60/074,725  
; PRIOR FILING DATE: 1998-02-13  
; PRIOR APPLICATION NUMBER: US 60/096,409  
; PRIOR FILING DATE: 1998-08-13  
; NUMBER OF SEQ ID NOS: 28208  
; SEQ ID NO 324  
; LENGTH: 2658  
; TYPE: DNA  
; ORGANISM: Candida albicans  
US-09-248-796A-324





;; PRIOR FILING DATE: 2000-10-20  
;; PRIOR APPLICATION NUMBER: 60/237,768  
;; PRIOR FILING DATE: 2000-10-03  
;; PRIOR APPLICATION NUMBER: 60/231,498  
;; PRIOR FILING DATE: 2000-09-08  
;; NUMBER OF SEQ ID NOS: 207012  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 14547  
;; LENGTH: 818128  
;; TYPE: DNA  
;; ORGANISM: Human  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; LOCATION: (1)...(818128)  
;; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-14547

Query Match 29.9%; Score 25.4; DB 3; Length 818128;  
Best Local Similarity 58.7%; Pred. No. 1.4e+02;  
Matches 44; Conservative 0; Mismatches 31; Indels 0; Gaps 0;  
  
Qy 3 GTGAAGTCATCTTCGCTGTTCTCTATTGAAACAAGCATTTCCCCAGGAAAGTTTCT 62  
Db 339125 GTGAATACCTTTTTCCTTGCTTCATTTAAATATATATGACCCCGTTTGTGTTTT 339184  
  
Qy 63 GGGTGCCCCCACTAAG 77  
Db 339185 GTTTTCCTTTTAG 339199

RESULT 14  
US-09-949-016-14548  
;; Sequence 14548, Application US/09949016  
;; Patent No. 6812339  
;; GENERAL INFORMATION:  
;; APPLICANT: VENTER, J. Craig et al.  
;; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
;; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
;; FILE REFERENCE: CL001307  
;; CURRENT APPLICATION NUMBER: US/09/949,016  
;; CURRENT FILING DATE: 2000-04-14  
;; PRIOR APPLICATION NUMBER: 60/241,755  
;; PRIOR FILING DATE: 2000-10-20  
;; PRIOR APPLICATION NUMBER: 60/237,768  
;; PRIOR FILING DATE: 2000-10-03  
;; PRIOR APPLICATION NUMBER: 60/231,498  
;; PRIOR FILING DATE: 2000-09-08  
;; NUMBER OF SEQ ID NOS: 207012  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 14548  
;; LENGTH: 818128  
;; TYPE: DNA  
;; ORGANISM: Human  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; LOCATION: (1)...(818128)  
;; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-14548

Query Match 29.9%; Score 25.4; DB 3; Length 818128;  
Best Local Similarity 58.7%; Pred. No. 1.4e+02;  
Matches 44; Conservative 0; Mismatches 31; Indels 0; Gaps 0;  
  
Qy 3 GTGAAGTCATCTTCGCTGTTCTCTATTGAAACAAGCATTTCCCCAGGAAAGTTTCT 62  
Db 339125 GTGAATACCTTTTTCCTTGCTTCATTTAAATATATATGACCCCGTTTGTGTTTT 339184  
  
Qy 63 GGGTGCCCCCACTAAG 77  
Db 339185 GTTTTCCTTTTAG 339199

RESULT 15

US-09-949-016-14549  
;; Sequence 14549, Application US/09949016  
;; Patent No. 6812339  
;; GENERAL INFORMATION:  
;; APPLICANT: VENTER, J. Craig et al.  
;; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
;; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
;; FILE REFERENCE: CL001307  
;; CURRENT APPLICATION NUMBER: US/09/949,016  
;; CURRENT FILING DATE: 2000-04-14  
;; PRIOR APPLICATION NUMBER: 60/241,755  
;; PRIOR FILING DATE: 2000-10-20  
;; PRIOR APPLICATION NUMBER: 60/237,768  
;; PRIOR FILING DATE: 2000-10-03  
;; PRIOR APPLICATION NUMBER: 60/231,498  
;; PRIOR FILING DATE: 2000-09-08  
;; NUMBER OF SEQ ID NOS: 207012  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 14549  
;; LENGTH: 818128  
;; TYPE: DNA  
;; ORGANISM: Human  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; LOCATION: (1)...(818128)  
;; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-14549

Query Match 29.9%; Score 25.4; DB 3; Length 818128;  
Best Local Similarity 58.7%; Pred. No. 1.4e+02;  
Matches 44; Conservative 0; Mismatches 31; Indels 0; Gaps 0;  
  
Qy 3 GTGAAGTCATCTTCGCTGTTCTCTATTGAAACAAGCATTTCCCCAGGAAAGTTTCT 62  
Db 339125 GTGAATACCTTTTTCCTTGCTTCATTTAAATATATATGACCCCGTTTGTGTTTT 339184  
  
Qy 63 GGGTGCCCCCACTAAG 77  
Db 339185 GTTTTCCTTTTAG 339199

Search completed: December 6, 2005, 11:47:38  
Job time : 47.0997 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 11:34:11 ; Search time 58.0228 Seconds  
(without alignments)  
455.919 Million cell updates/sec

Title: US-09-980-046B-10  
Perfect score: 85  
Sequence: 1 cggtagagtgcatcttctgc.....tgccccactaagtagaaaaa 85

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3289935 seqs, 155610033 residues

Total number of hits satisfying chosen parameters: 6579870

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA New: \*  
1: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	26.8	31.5	3934	US-10-750-185-49893 Sequence 49893, A
2	26.4	31.1	1383	US-10-750-185-63714 Sequence 63714, A
3	25.6	30.1	3331	US-10-750-185-46975 Sequence 46975, A
4	25.4	29.9	1846	US-10-750-185-44982 Sequence 44982, A
5	25.4	29.6	2786	US-10-750-185-58054 Sequence 58054, A
6	25.2	29.6	1846	US-10-750-185-31717 Sequence 31717, A
7	25	29.4	2518	US-10-750-185-36838 Sequence 36838, A
8	24.8	29.2	4341	US-10-750-185-43918 Sequence 43918, A
9	24.6	28.9	1364	US-10-750-185-43944 Sequence 43944, A
10	24.2	28.5	1535	US-10-750-185-46482 Sequence 46482, A
11	24	28.2	905	US-10-750-185-47332 Sequence 47332, A
12	24	28.2	1795	US-10-750-185-37876 Sequence 37876, A
13	23.8	28.0	600	US-10-750-185-1830 Sequence 1830, Ap
14	23.8	28.0	600	US-10-750-185-1860 Sequence 1860, Ap
15	23.8	28.0	1166	US-10-750-185-59439 Sequence 59439, A
16	23.8	28.0	1411	US-10-750-185-62589 Sequence 62589, A
17	23.6	27.8	1726	US-10-750-185-57813 Sequence 57813, A
18	23.4	27.5	1380	US-10-750-185-62655 Sequence 62655, A
19	23.2	27.3	1634	US-10-750-185-64591 Sequence 64591, A
20	23.2	27.3	2541	US-10-750-185-27749 Sequence 27749, A
21	23.2	27.3	2558	US-10-750-185-52868 Sequence 52868, A
22	23	27.1	1509	US-10-750-185-50494 Sequence 50494, A
23	23	27.1	2628	US-10-750-185-53005 Sequence 53005, A

ALIGNMENTS

RESULT 1

US-10-750-185-49893  
; Sequence 49893, Application US/10750185  
; Publication No. US20050260603A1  
; GENERAL INFORMATION:  
; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MM1100-2  
; CURRENT APPLICATION NUMBER: US/10750185  
; CURRENT FILING DATE: 2003-12-31  
; PRIOR APPLICATION NUMBER: US 60/437,482  
; PRIOR FILING DATE: 2002-12-31  
; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 49893  
; LENGTH: 3934  
; TYPE: DNA  
; ORGANISM: Bovine 19866881023423  
US-10-750-185-49893

Query Match 31.5%; Score 26.8; DB 6; Length 3934;  
Best Local Similarity 59.0%; Pred. No. 0.96;  
Matches 46; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 4 TGAGTGCATCTTCGGTCTCTATTTCGAACAAGCATTTCCCCAGGAGTTCTG 63  
DB 2715 TTAAGAGGATCATCTGGCTGTTTCTCTATTTCCTAAGCATTTAAGCATTTCAAGTGTGCG 2774  
QY 64 GGTCGCCCACTAAGTAGA 81  
DB 2775 CTTGACCTCGAAGTATA 2792

RESULT 2

US-10-750-185-63714/c  
; Sequence 63714, Application US/10750185  
; Publication No. US20050260603A1  
; GENERAL INFORMATION:  
; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.

```
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 63714
; LENGTH: 1383
; TYPE: DNA
; ORGANISM: Bovine 19866881014087
US-10-750-185-63714

Query Match      31.1%; Score 26.4; DB 6; Length 1383;
Best Local Similarity 59.2%; Pred. No. 0.9;
Matches 45; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

Qy 4 TGAAGTGCATCTTCTGGGTGTTCTATTGAACAAGCATTTCCCCAGGGAAGTTTCTG 63
Db 303 TGCAGTCGAGAGACAGCGGTTTCATTCCTGTGTCAGAGAGTCCCCAGAGAGGAGTG 244

Qy 64 GGTCCCCCACTAAGTA 79
Db 243 GCTCCCCCAATCCAGTA 228
```

```
RESULT 3
US-10-750-185-46975
; Sequence 46975, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 46975
; LENGTH: 3331
; TYPE: DNA
; ORGANISM: Bovine 19866881406138
US-10-750-185-46975

Query Match      30.1%; Score 25.6; DB 6; Length 3331;
Best Local Similarity 70.8%; Pred. No. 2.6;
Matches 34; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 1 CGGTGAAGTGCATCTTCTGGGTGTTCTATTGAACAAGCATTTCCC 48
Db 1778 CTGTAAGAGAAAGAACTTCTCCTCATCTACTATTGTAATGATTTC 1825
```

```
RESULT 4
US-10-750-185-44982/c
; Sequence 44982, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
```

```
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 44982
; LENGTH: 1846
; TYPE: DNA
; ORGANISM: Bovine 19866880725399
US-10-750-185-44982
```

```
Query Match      29.9%; Score 25.4; DB 6; Length 1846;
Best Local Similarity 68.6%; Pred. No. 2.5;
Matches 35; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 33 TGAACAAGCATTTCCCCAGGGAAGTTTCTGGGTGCCCACTAAGTAGAAA 83
Db 1347 TGAACAACATATGCGCATAGGAAGGCTGCTGGATGCTCTACTATTGAGAAA 1297
```

```
RESULT 5
US-10-750-185-58054/c
; Sequence 58054, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 58054
; LENGTH: 2786
; TYPE: DNA
; ORGANISM: Bovine 19866880580668
US-10-750-185-58054
```

```
Query Match      29.9%; Score 25.4; DB 6; Length 2786;
Best Local Similarity 61.2%; Pred. No. 2.9;
Matches 41; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

Qy 19 GCGTGTCTCTATTGAAACAAGCATTTCCCAAGGGAAGTTTCTGGGTGCCCACTAAGT 78
Db 2267 GGGCATTTCTCATATGTACAAAGATTCTCTAGAAAAAATTTATTAGAAACCCCTCTAAGT 2208

Qy 79 AGAAAAA 85
Db 2207 AAAATAA 2201
```

```
RESULT 6
US-10-750-185-31717/c
; Sequence 31717, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
```

```
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 31717
; LENGTH: 1846
; TYPE: DNA
; ORGANISM: Bovine 19866880465309
US-10-750-185-31717

Query Match          29.6%; Score 25.2; DB 6; Length 1846;
Best Local Similarity 62.9%; Pred. No. 2.9;
Matches 39; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 23 GTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTTCTGGTGCCCACTTAAGTAGAA 82
Db 1680 GATTTCATATGACACAAACCCTTCAAGATGATGTTGGGGGCATCCACACGAGAAA 1621

QY 83 AA 84
Db 1620 AA 1619
```

```
RESULT 7
US-10-750-185-36838/c
; Sequence 36838, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 36838
; LENGTH: 2518
; TYPE: DNA
; ORGANISM: Bovine 19866880875187
US-10-750-185-36838

Query Match          29.4%; Score 25; DB 6; Length 2518;
Best Local Similarity 58.9%; Pred. No. 4;
Matches 43; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 2 GGTGAAGTCATCTTCTGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTTC 61
Db 2434 GCTGCAGTGTAGATTCTGGGTCTCTGTAGGACCCAGGTTTGATCCCTGGGTCGAGTC 2375

QY 62 TGGTGCCCACT 74
Db 2374 AGGAAGATCCCT 2362
```

```
RESULT 8
US-10-750-185-44918
; Sequence 44918, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 44918
; LENGTH: 4341
; TYPE: DNA
; ORGANISM: Bovine 19866880547884
US-10-750-185-44918

Query Match          29.2%; Score 24.8; DB 6; Length 4341;
Best Local Similarity 60.3%; Pred. No. 5.9;
Matches 41; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 12 ATCTTCTGCGTGTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTTCTGGGTGCCCC 71
Db 3179 ATCTAATGTCATTCATGATTTTAAAGCAACATTCACGGGGAATTCATGCTGGTGCCC 3238

QY 72 ACTAAGTA 79
Db 3239 AGTGCCTA 3246
```

```
RESULT 9
US-10-750-185-43944/c
; Sequence 43944, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 43944
; LENGTH: 1364
; TYPE: DNA
; ORGANISM: Bovine 19866881010494
US-10-750-185-43944

Query Match          28.9%; Score 24.6; DB 6; Length 1364;
Best Local Similarity 57.0%; Pred. No. 4.4;
Matches 45; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 6 AAGTCATCTTCTGGGTGTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTTCTGGG 65
Db 851 AATTGGAACCTATTTGTAGGTTCTGTTTGGGAAACATGCTTTAGTTCTGTTTGGGA 792
```



Matches 46; Conservative 0; Mismatches 37; Indels 0; Gaps 0;  
Qy 2 GGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTTC 61  
Db 187 GCTTATGTGGTCTTTTGTGTTGTTTTCGGTCATGCCCTATAGCATGCAAGATCT 246  
Qy 62 TGGGTGCCCCACTAAGTAGAGAAA 84  
Db 247 TAGTTCCCAACCAAGGAGTGAA 269

RESULT 14

US-10-750-185-1960  
; Sequence 1960, Application US/10750185  
; Publication No. US20050260603A1  
; GENERAL INFORMATION:  
; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MM1100-2  
; CURRENT APPLICATION NUMBER: US/10/750,185  
; CURRENT FILING DATE: 2003-12-31  
; PRIOR APPLICATION NUMBER: US 60/437,482  
; PRIOR FILING DATE: 2002-12-31  
; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1960  
; LENGTH: 600  
; TYPE: DNA  
; ORGANISM: Bovine MMBT21455  
US-10-750-185-1960

Query Match 28.0%; Score 23.8; DB 6; Length 600;  
Best Local Similarity 55.4%; Pred. No. 6.5;  
Matches 46; Conservative 0; Mismatches 37; Indels 0; Gaps 0;  
Qy 2 GGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTTC 61  
Db 97 GCTTATGTGGTCTTTTGTGTTGTTTTCGGTCATGCCCTATAGCATGCAAGATCT 156  
Qy 62 TGGGTGCCCCACTAAGTAGAGAAA 84  
Db 157 TAGTTCCCAACCAAGGAGTGAA 179

RESULT 15

US-10-750-185-59439  
; Sequence 59439, Application US/10750185  
; Publication No. US20050260603A1  
; GENERAL INFORMATION:  
; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MM1100-2  
; CURRENT APPLICATION NUMBER: US/10/750,185  
; CURRENT FILING DATE: 2003-12-31  
; PRIOR APPLICATION NUMBER: US 60/437,482  
; PRIOR FILING DATE: 2002-12-31  
; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 59439  
; LENGTH: 1166  
; TYPE: DNA

; ORGANISM: Bovine 19866880972937  
US-10-750-185-59439  
Query Match 28.0%; Score 23.8; DB 6; Length 1166;  
Best Local Similarity 62.7%; Pred. No. 8.5;  
Matches 37; Conservative 0; Mismatches 22; Indels 0; Gaps 0;  
Qy 25 TCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTTCTGGGTGCCCCACTAAGTAGAAA 83  
Db 350 TATCTTTTGTCCAGCCATCTGCTTTGAAACAAGTTTCTGTCCAGCCTCTGGAAGTAGTAA 408

Search completed: December 6, 2005, 23:16:55  
Job time : 58.0228 secs

**This Page Blank (uspto)**